

PHYSIOLOGICAL DEMANDS OF ELITE SOCCER REFEREEING: NEEDS ANALYSIS AND APPLICATIONS TO TRAINING AND MONITORING

MATTHEW WESTON

A thesis submitted in partial fulfilment of the requirements of the
University of Teesside for the degree of Doctor of Philosophy

September 2009

TABLE OF CONTENTS

Abstract	i
Acknowledgments	ii
Chapter 1: Preface to the PhD via completed work	
1.1 Introduction to the field of study	1
Chapter 2: Training Performance	
2.1 The impact of specific high-intensity training sessions on football referees' fitness levels.....	12
Chapter 3: Match Demands	
3.1 The effect of match standard and referee experience upon the objective and subjective match workload of English Premier League referees.....	21
3.2 Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates	29
Chapter 4: Ageing and Performance	
4.1 Age-related effects on fitness performance in elite-level soccer referees	38
4.2 Ageing and physical match performance in English Premier League soccer referees	45
Chapter 5: Fitness and Match Performance	
5.1 Relationship between field tests and physical match performance in elite soccer referees	51
Chapter 6: Summary and Implications for future research	60
APPENDIX	99

ABSTRACT

The following thesis investigates contemporary issues within the applied physiology of soccer refereeing. 1) Training performance. The impact of a high intensity training regime was examined in a group of elite-level soccer referees. Following a 16-month training period the referees' performance on the YoYo Intermittent Recovery Test (level 1) improved by 46.5%. 2) Match Demands. The effect of match standard and referee experience upon the objective and subjective match loads of referees was investigated. Match heart rates (HR) and ratings of perceived exertion (RPE) were both related to standard of competition, with the match loads being higher on the higher standard of competition. Referee experience had no effect upon the referees' match responses. Using a semi-automatic, video match analysis system the referees' match activities and factors affecting these activities were also examined. Physical performances were related in part to the physical performances of the players; whilst the distances covered during the first half were related to second half coverage. 3) Ageing and performance. The effect of ageing upon referees' fitness levels and physical match performances was addressed. Regression analysis revealed a trend towards an age-related reduction in physical fitness, as determined by the referees' fitness tests. Match activity analysis demonstrated a clear age-related decline in physical match performance, although this decline did not impair the referees' ability to keep up with play. 4) Fitness and match performance. The validity of the FIFA referees' fitness tests was examined. Interval test HR load was significantly correlated to the referees' match coverage, both total distance and high intensity running. Sprint test scores also demonstrated a significant relationship with the referees' match sprinting distances. However, given the strength of the relationships only the sprint test showed appropriate construct validity for the physical assessment of soccer referees.

ACKNOWLEDGEMENTS

The manuscripts presented in this thesis represent almost a decade of work, spanning my initial involvement with elite-level referees, though to work with UEFA and FIFA and finally as sports scientist with the Professional Game Match Officials Limited (PGMOL). Therefore, I would like to thank all of the co-authors whose valuable contributions to each study has been greatly appreciated, in particular the guidance of Carlo Castagna and Franco Impellizzeri. I would also like to acknowledge the help and support from all of the referees involved in the research, along with the PGMOL and ProzoneTM, in particular Keith Hackett and Martin Bland.

My sincere thanks go to Professor Alan Batterham for acting as my thesis advisor and also for his expert advice, guidance and supervision over the past few months. Thanks also to Matt Portas for getting the process at Teesside University started.

I would like to thank all my friends and family for their support of the years. Finally, I save the biggest thank you of them all to Jessica for her unwavering love and support through each step, each project, each season, and each move.

1.0 Preface to the PhD via completed work

1.1 Introduction to the field of study

The scientific papers presented in this thesis focus upon the physical performances of soccer referees, during competitive matches and also during training. Soccer (or association football) is the most popular sport in the world (Bangsbo, 1994) with each soccer match being controlled by a referee (Stølen, Chamari, Castagna & Wisløff, 2005). Soccer referees are charged with the responsibility for implementing the rules of the game and guaranteeing that players abide by its regulations (Reilly & Gregson, 2006). At elite-level, referees have been demonstrated to make 137 (range 104-162) observable decisions per match, which given an effective playing time of 51 min, results in a top-class referee making 3-4 decisions per minute (Helsen & Bultynck, 2004). Consequently, the soccer referee is required to keep up with play at all times and ensure optimal viewing positions in order to make the correct judgements.

Early research into the match activity profiles of soccer referees demonstrated that officiating at elite-level places high physiological demands on the referee. In a study examining referees' work rate and heart rates during association football matches in England, Catterall, Reilly, Atkinson & Coldwells (1993) reported a mean (range) distance covered by the referees during a 90-min match of 9348m (7977 – 10187m). This distance was consistent with the 9408 ± 838 m reported for Australian soccer referees (Johnston & McNaughton, 1994). More recently, results obtained from Krstrup & Bangsbo (2001) and D'Ottavio & Castagna

(2001) have demonstrated an increase in referees' match work rate over the past decade. Krustup & Bangsbo (2001) reported a mean (range) match distance coverage of 10070m (9200 – 11490m) with standing and walking accounting for 63.2% of total match time, low-intensity running 30.2% and high-intensity running accounting for the remainder 6.6%. In terms of actual distances covered, D'Ottavio & Castagna (2001) reported a mean distance over the course of an average 95-min match of 11376 ± 1604 m, with walking constituting 889 ± 94 m, 4174 ± 616 m of low-intensity running ($<13\text{km.h}^{-1}$), 2585 ± 489 m of medium-intensity running ($13.1 - 18.0\text{km.h}^{-1}$) and a total of 1556 ± 493 m of high-intensity running ($18.1 - 24\text{km.h}^{-1}$). Running at maximal speeds ($>24\text{km.h}^{-1}$) accounted for 608 ± 366 m and non-orthodox directional modes, i.e., sideways and backwards running, for 181 ± 117 m and 1315 ± 569 m, respectively.

Heart rate recordings have also demonstrated that soccer refereeing at an elite-level places a high physiological demand upon referees (Johnston & McNaughton, 1994). Catterall *et al.* (1993) reported a mean heart rate for all referees over the course of 13 matches of $165 \pm 8 \text{ b.min}^{-1}$, which corresponded to 95% of the referees' age-predicted maximal heart rate (HR_{max}). Johnston & McNaughton (1994) and D'Ottavio & Castagna (2001), also using an age-predicted HR_{max} , reported similarly high match heart rates.

However, the age-predicted formula ($220 - \text{age}$) for HR_{max} used within the study's methodology has been reported to overestimate HR_{max} in men and women less than 40 years and underestimate HR_{max} for those older than 40 years (Gellish *et al.*, 2007). Consequently, lower mean match heart rates of $85\%\text{HR}_{\text{max}}$

(74-91%) were reported by Krstrup & Bangsbo (2001) for Danish referees whose HR_{max} was determined directly by a progressive treadmill test protocol. Helsen & Bultynck (2004) also reported elite-level referees' mean match heart rates to be $85 \pm 5\%HR_{max}$, with HR_{max} being taken as the peak 5-s value recorded on either a laboratory test, training session or match.

Further examinations into the physiology of soccer refereeing have been provided through oxygen uptake and blood lactate assessments, which provide an indication of the aerobic and anaerobic involvement during matches. Krstrup & Bangsbo (2001) estimated, through the use of a laboratory measured heart rate- VO_2 relationship, that the aerobic involvement during competitive Danish soccer matches was 81% (73 – 88%) of the referees' VO_{2max} . To obtain a more direct measure of the aerobic involvement of soccer refereeing D'Ottavio & Castagna (2002) measured elite Italian soccer referees' match VO_2 responses using a lightweight, portable gas analyser. The data were collected on non-competitive matches over the duration of one half only (45 min). The authors reported that the referees attained on average 68% of their individual VO_{2max} . Given that the Italian referees covered the same distances and recorded the same heart rates as on competitive matches, Castagna, Abt & D'Ottavio (2007) accounted this difference in results to the different methodologies employed to assess VO_2 as the use of the treadmill heart rate- VO_2 may overestimate the actual match VO_2 .

Blood lactate measurements can provide a gross indication of the involvement of anaerobic glycogenolysis (Castagna *et al.*, 2007) and Krstrup & Bangsbo (2001) reported concentrations of 4.8 (2.0 – 9.8 mmol/L) and 5.1mmol/L (2.3 – 14.0

mmol/L) at the end of the first and second halves, respectively for elite Danish soccer referees on competitive matches. D'Ottavio & Castagna (2002) reported values in the region of 7mmol/L for soccer referees officiating on non-competitive matches. However, in both studies large intra- and inter-individual differences were observed and it is important to note that lactate measures during actual match play are influenced largely by the activity pattern in the 5-min period that precedes the blood sampling (Stølen *et al.*, 2005).

Sports science is a multi-disciplinary field concerned with the understanding and enhancement of human sporting performance; a scientific process used to guide the practice of sport with the ultimate aim of improving sports performance (Bishop, 2008). The physiology discipline of the sports sciences involves fitness testing, training prescription (and monitoring), match analysis and athlete / coach education. However, for the scientist to be truly effective they must first develop both theoretical and practical knowledge of their sport.

Sport and exercise scientists engaged in soccer are interested in a multitude of factors that determine the performance of a player as well as the related underlying phenomena that explain how each factor influences that performance (Drust, Atkinson & Reilly, 2007). However, within the scientific literature there appear to be many articles that have arisen from convenient access to a data set rather than being driven by a problem to be solved (Bishop, 2008). Consequently the scope of each of the six studies presented in this thesis was to move beyond this by investigating real-world problems with regards to the applied physiology of soccer refereeing and in turn produce practically relevant applied research that

examined some of the key issues and problems that arose from day-to-day work with elite-level referees. For example, what are the physiological demands of soccer refereeing at the elite-level in England, what factors influence these demands and what is the trainability of an 'older' population utilising modern training techniques?

In order to build practice from evidence and theory it is important to not start with the theory and look for problems on which to test them, but to start with problems and look for theories to help solve them (Green, 2006). Therefore, utilising such an approach the intended outcome of this programme of work was to produce practice-based research findings that are usable and relevant with regards to the physical performance of soccer referees. Ultimately this process may help direct a more evidence-based framework for the application of sports science; making the practice more theory-based.

Seiler & Kjerland (2006) reported that experimental studies are extremely difficult to perform on high-level athletes because neither the athletes nor their coaches wish to suddenly alter methods they have developed over perhaps years of coaching; a common issue reported by studies where the subjects have been elite-level athletes. Therefore, the research models contained within this thesis lie mostly towards the applied end of the basic-applied research continuum presented by Atkinson & Nevill (2001) and expanded upon by Drust *et al.* (2007). Specifically, the studies examined factors affecting the referees' physical performances in realistic situations (performance tests and matches) and evaluated if one variable (e.g., age, experience, fitness, standard of competition

etc.) made a difference to another during competitive matches and fitness tests. Given that real-life refereeing situations were under investigation the studies would possess high levels of ecological validity.

Traditionally soccer referees' fitness training was centred on long, steady state aerobic exercises (Krustrup & Bangsbo, 2001) with very little sprinting and high-intensity running. Whilst such exercises may have developed an appropriate level of fitness required to pass the annual referee fitness test they may not have prepared the referees in a manner specific to the demands of their games. Therefore, with the advent of full-time sports science support to the referees, more modern training methods were introduced into the referees' routines. Therefore, STUDY I examined the heart rate responses during specific intensive intermittent training sessions, whilst also examining the impact of such training methods upon the referees' fitness levels.

Unlike full-time soccer players, the English Premier League referees only have group training sessions every other week. Therefore, the referees' training sessions have to be monitored from afar and since 2001 all of the referees have their objective (heart rate) and subjective (ratings of perceived exertion; RPE) ratings of exercise intensity monitored on all training sessions and matches. This permits the intensity of all sessions to be evaluated in a remote way. The referees record their RPE score, using Borg's CR10 scale (Borg, Hassmen & Lagerstrom, 1987), 30 min after the session ends in order to obtain a global intensity rating for the entire session (Foster, 1998). Therefore, a particularly easy or difficult bout of exercise towards the end would not dominate the rating. This scale has

been reported to correlate significantly with objective indices of exercise training such as heart rate and blood lactate (Foster *et al.*, 2001). Heart rate was recorded every 5-s via a short-range telemetry system (Polar S610, Kemple, Finland).

The collection of heart rate and RPE data on competitive matches during the 2002/2003 English soccer season (STUDY II) utilising a controlled research design enabled a greater understanding of the physical demands of elite-level soccer refereeing and also the effect of match standard as data were not only collected on Premier League matches but also on Football League matches, a lower standard of competition.

From 2003 onwards the referees' match activity profiles have been collected at selected Premier and Football League grounds. This data is collected using the Prozone[®] system which uses eight cameras to record the positional data of every player and official, every 1/10th of a second during a match. Recent findings have demonstrated that the Prozone[®] match-analysis system provides valid and reliable analyses of movement patterns of footballers during matchplay (Di Salvo, Collins, McNeill & Cardinale, 2006; DiSalvo, Gregson, Atkinson, Tordoff & Drust, 2009). With the collection of the referees' match activity profiles it was possible to examine in more detail their physical match performances and some of the factors impacting upon these performances as it was clear from previous studies (Krustrup & Bangsbo, 2001; D'Ottavio & Castagna 2001; Castagna & Abt, 2003) that the physical demands imposed upon referees vary significantly from match to match. Therefore, STUDY III was performed in an attempt to gain

a greater understanding into the physical match performances of elite-level English soccer referees.

An important aspect of soccer refereeing is the existing age difference between soccer players and soccer referees (Stølen *et al.*, 2005). The referee is required to keep up with play at all times despite occupying an age bracket on average 10 to 15 years older than their playing counterparts (Catterall *et al.*, 1993). However, ageing is a multi-factoral process than can potentially have a profound effect upon the physiology aspect of human performance (Stamford, 1988). With this in mind the effect of age upon fitness levels in elite-level soccer referees was examined in STUDY IV, with STUDY V examining the effect of age upon the physical match performances and match physiological load of elite-level soccer referees.

With regards to referee selection, a greater priority is placed upon fitness levels as opposed to physical match performances as national and international soccer refereeing associations use the Fédération Internationale de Football Association (FIFA) fitness tests as selection criteria for competitions. Nonetheless fitness tests need to accurately reflect a referee's match fitness. Therefore, to determine whether the FIFA established referees' fitness tests actually reflect a referee's match-related physical capacity STUDY VI examined the relationship between fitness test performances with physical match performance.

References

- Atkinson, G. & Nevill, A.M. (2001) Select issues in the design and analysis of sports performance research. *Journal of Sports Sciences*, 19, 811-827.
- Bangsbo, J. (1994) The physiology of soccer: with special reference to intense intermittent exercise. *Acta Physiologica Scandinavica*, 15 S619, 1-156.
- Bishop, D. (2008) An applied research model for the sports sciences. *Sports Medicine*, 38, 253-263.
- Borg, G., Hassmen, P. & Lagerstrom, M. (1987) Perceived exertion related to heart rate and blood lactate during arm and leg exercise. *European Journal of Applied Physiology and Occupational Physiology*, 56, 679-685.
- Castagna, C. & Abt, G. (2003) Intermatch variation of match activity in elite Italian soccer referees. *Journal of Strength and Conditioning Research*, 17, 388-392.
- Castagna, C., Abt, G. & D'Ottavio, S. (2007) Physiological aspects of soccer refereeing performance and training. *Sports Medicine*, 37, 625-646.
- Catterall, C., Reilly, T., Atkinson G. & Coldwells, A. (1993) Analysis of the work rates and heart rates of association football referees. *British Journal of Sports Medicine*, 27, 193-196.
- D'Ottavio, S. & Castagna, C. (2001) Physiological load imposed on elite soccer referees during actual match play. *Journal of Sports Medicine and Physical Fitness*, 41, 27-32.
- D'Ottavio, S. & Castagna, C. (2002) Physiological aspects of soccer refereeing. In *Science and Football IV* (Ed. Spinks, W., Reilly, T. and Murphy, A.) London:Routledge, 144-150.

- Di Salvo, V., Collins, A., McNeill, B. & Cardinale, M. (2006) Validation of Prozone: A new video-based performance analysis system. *International Journal of Performance Analysis in Sport*, 6, 108-19.
- Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P. & Drust, B. (2009) Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine*, 30, 205-212
- Drust, B., Atkinson, G. & Reilly, T. (2007) Future perspectives in the evaluation of the physiological demands of soccer. *Sports Medicine*, 37, 783-805.
- Foster, C. (1998) Monitoring training in athletes with reference to overtraining syndrome. *Medicine and Science in Sports and Exercises*, 30, 1164-1168.
- Foster, C., Florhaug, J.A., Franklin, J., Gottschall, L., Hrovatin, L.A., Parker, S., *et al.* (2001) A new approach to monitoring exercise training. *Journal of Strength and Conditioning Research*, 15, 109-15.
- Gellish, R.L., Goslin, B.R., Olsen, R.E., Macdonald, A., Russi, G.D. & Moudgil, V.K. (2007) Longitudinal modelling of the relationship between age and maximal heart rate. *Medicine and Science in Sports and Exercise*, 27, 822-829.
- Green, L.W. (2006) Public health asks of systems sciences: to advance our evidence-based practice, can you help us get more practice-based evidence? *American Journal of Public Health*, 96, 406-409.
- Helsen, W. & Bultynck, J.B. (2004) Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, 22, 179-189.
- Johnston, L. & McNaughton, L. (1994) The physiological requirements of soccer refereeing. *Australian Journal of Science and Medicine in Sport*, 26, 67-72.

- Krustrup, P. & Bangsbo, J. (2001) Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. *Journal of Sports Sciences*, 19, 881-891.
- Reilly, T. & Gregson, W. (2006) Special populations: The referee and assistant referee. *Journal of Sports Sciences*, 24, 795-801.
- Seiler, K.S. & Kjerland, G. ø. (2006) Quantifying training intensity distribution in elite endurance athletes: is there evidence for an “optimal” distribution? *Scandinavian Journal of Medicine and Science in Sports*, 16, 49-56.
- Stamford, B.A. (1988) Exercise and the Elderly. *Exercise and Sports Science Reviews*, 16, 341-379.
- Stølen, T., Chamari, K., Castagna, C. & Wisløff, U. (2005) Physiology of soccer: an update. *Sports Medicine*, 35, 501-536.

CHAPTER 2.0

TRAINING PERFORMANCE

2.1

STUDY I

Weston, M., Helsen, W., MacMahon, C. & Kirkendall, D. (2004) The impact of specific high intensity training sessions on football referees' fitness levels. *American Journal of Sports Medicine*, 32, S54-S61.

The Impact of Specific High-Intensity Training Sessions on Football Referees' Fitness Levels

Matthew Weston,* BSc(Hons), MSc, Werner Helsen,*† PhD, Clare MacMahon,‡ BSc, MSc, and Don Kirkendall,§ MD, PhD

*From the *Department of Kinesiology, Katholieke Universitat Leuven, Belgium, the †Department of Kinesiology, McMaster University, Hamilton, Ontario, Canada, and the §Department of Orthopaedics, University of North Carolina*

Background: In comparison to the amount of literature that has examined the match demands of football refereeing, there has been little attempt to assess the impact of high-intensity training.

Purpose: The main goals were to get a better understanding of the long-term effect of specific intermittent training.

Study Design: The authors examined the cardiovascular strain of specific high-intensity training sessions and also their impact on referees' fitness levels.

Methods: To examine the physical workload during intensive intermittent training sessions, heart rates were recorded and analyzed relative to the referees' maximum heart rate (HR_{max}). To assess the referees' fitness levels, the Yo-Yo intermittent recovery test was used.

Results: Both the pitch- and track-training sessions were successful in imposing an appropriate high intensity load on the referees, at $86.4 \pm 2.9\%$ and $88.2 \pm 2.4\%$ HR_{max} , respectively. Following 16 months of intermittent high-intensity training, referees improved their performance on the Yo-Yo intermittent recovery test by 46.5%, to a level that is comparable with professional players.

Conclusions: As match officials are subjected to a high physical load during matches, they should follow structured weekly training plans that have an emphasis on intensive, intermittent training sessions.

Keywords: intermittent training; Yo-Yo test; football refereeing

The physiological characteristics of football referees have received an increasing amount of focus in the scientific literature over the past decade. Studies have examined referee profiles, both physiological and anthropometric,¹⁰ as well as the movement patterns and physiological load experienced during actual match play.^{6,7,10-12,21} Match analysis by D'Ottavio and Castagna⁷ demonstrated that referees cover a mean \pm SD distance, over the course of an average 95-minute match, of 11469 ± 983 m, with walking constituting 957 m, 4174 m of low-intensity running, 2585 m of medium-intensity running, and a total of 1556 m of high-intensity running. Running at maximal speed accounted for 608 m, and nonorthodox directional modes—that is, sideways and backward running—accounted for 181 m and 1315 m, respectively.⁷

† Address correspondence and reprint requests to Dr. Werner Helsen, Katholieke Universiteit Leuven, Tervuursevest 101, Leuven 3001, Belgium.

The American Journal of Sports Medicine, Vol. 32, No. 1 Suppl.
DOI: 10.1177/0363546503261421
© 2004 American Orthopaedic Society for Sports Medicine

Heart rate recordings collected during matches demonstrated that the mean cardiovascular strain imposed on referees during match is approximately 85% of maximal heart rate (HR_{max}).^{9,11,21} However, as referees have been reported to spend approximately 75% of their total match time either standing, walking, or jogging,¹¹ the amount of high-intensity activity, performed at speeds of more than 15 km/hour during a match, is a better indicator of the physical demands of matches.^{3,11,12}

Given the high physical load imposed on top-class referees during actual match play,^{7,9,11} fitness levels need to be sufficient enough for the referees to be able to cope with the demands of their games through keeping up with play at all times and ensuring optimal viewing positions. This in itself is a challenge. However, when combined with the fact that referees are on average 10 to 15 years older than their playing counterparts^{6,7,9} and aging has a negative effect on fitness levels,^{20,23} referees have to work extremely hard in training to ensure that they attain, and maintain, an appropriate level of fitness. Also, in most countries referees

still work full-time, and their physical training sessions often have to be arranged around work commitments. Therefore, to ensure referees can attain an optimal level of match fitness, emphasis within their fitness preparation programs has to be firmly placed on quality structured training sessions that provide an appropriate training stimulus to enable the attainment of such fitness levels, especially with training time being at such a premium.

To overcome this problem, the governing bodies of European and world football, the Union des Associations Européennes de Football (UEFA) and the Fédération Internationale de Football Association (FIFA), have moved to further professionalize the football referees' environment over recent seasons. After all, the game has become more athletic and faster in recent years,²² and the referee or match official can no longer remain outside these developments. For referees, this is demonstrated by the fact that Catterall et al.⁶ reported a mean \pm SD distance covered by the referees during matches in 1993 of 9438 ± 707 m, whereas some 8 years later it was reported to be 11376 ± 1600 m.⁷ Some suggest that referees should be provided with high-tech tools, for example, to detect if the ball has fully crossed the goal line or to reexamine a tricky piece of action using video replays of the ongoing game. However, others question these kinds of technical assistance, arguing their rejection in terms of the need to preserve the human input and therefore imperfect nature of the game. The point of view we take in this research study is to provide the referees with contemporary, match-specific physical preparation sessions, which replicate the physical demands of football refereeing.

Despite the amount of literature that has profiled the match demands of football refereeing, to our knowledge there is no published literature in football, or any other team sports, that has examined the long-term effect of structured training programs and, more specifically, high-intensity training on referees' fitness levels. Given that the performance of referees during matches has been reported to be equal to that of midfield players and even greater than that of other outfield players,¹⁷ it is surprising that the Krustup and Bangsbo¹¹ study has been the only attempt to date to examine the impact of training on referees' fitness levels. The characteristics of their training sessions were based around high-intensity running at 90% HR_{max} with a work-rest ratio of 2:1. The sessions were intermittent, with either long intervals (4 or 8 minutes) or short intervals (30 seconds or 1 minute). However, due to the short-term nature of the training study, the long-term effects of such high-intensity running on referees' fitness levels remain unexamined. Also, despite reporting the intensity at which the high-intensity running was performed, the study failed to report any of the actual training heart rate data.

With this in mind, the aim of this study was to present the type and exact intensity of high-intensity intermittent training sessions that are being prescribed to top-class UEFA and FIFA referees as part of their weekly training regime and how these types of sessions affect referees' heart rates during the training sessions. Also, the long-term effectiveness of such training sessions with regard to

referees' fitness levels, as determined by the Yo-Yo intermittent recovery test, was also examined.

METHOD

Participants

The participants for this study consisted of 18 Belgian referees who were separated into two groups according to ability. The first group consisted of 7 Belgian referees (mean age 37.8 ± 4.12 years) who were also listed FIFA international referees. The referees had 13.7 ± 4.15 years of experience refereeing within their own country and 4.4 ± 4.16 years of experience refereeing internationally. The second group consisted of 11 Belgian elite referees (mean age 39.28 ± 4.21 years). These referees were affiliated with the Belgian football federation (KBVB) and only refereed nationally in the first and second Belgian leagues. The data were collected as part of the ongoing physiological support provided to the Belgian referees' committee.

Physical Workload During Training Sessions

Match-specific weekly training plans for the referees were introduced from the beginning of the 2001/2002 Belgian football season. This involved the referees receiving a training plan by e-mail for each week of their competitive season (and also preseason), with the plan being based around four sessions per week: recovery, high intensity, speed endurance, and speed training. The focus of this study is on the specific high-intensity training sessions. However, it would be impractical to present all the high-intensity sessions that the referees have been prescribed over the past two seasons since their training regimes became match specific in terms of the physical demands of football refereeing. Therefore, the sessions outlined in this study are actual examples of the high-intensity training sessions included in the referees' weekly training plans.

The referees' heart rates during the training sessions were recorded via short-range telemetry using Polar S610 watches (Polar, Kempele, Finland) with the data being recorded every 5 seconds. Following each group training session, the data were downloaded onto a computer for further analysis. Analysis was performed using the Polar Precision Performance software, version 3.0 (Polar, Kempele, Finland).

There are two types of sessions presented, depending on if they were performed on a football pitch (exercises 1 and 2) (Fig. 1) or on a 400-m running track (exercises 3 and 4). For the purpose of this study, only the heart rate data recorded during the high-intensity and active-recovery running within sets are reported. The heart rate data during the recovery in between sets were not included in the analysis.

For exercise 1, all referees perform the exercise in pairs. On a signal from the coach, one referee (A) from each pair starts running at 90% HR_{max} clockwise around the diagonal carrying a vest. The partner referee (B) from each pair walks on the outside part of the diagonal running area.

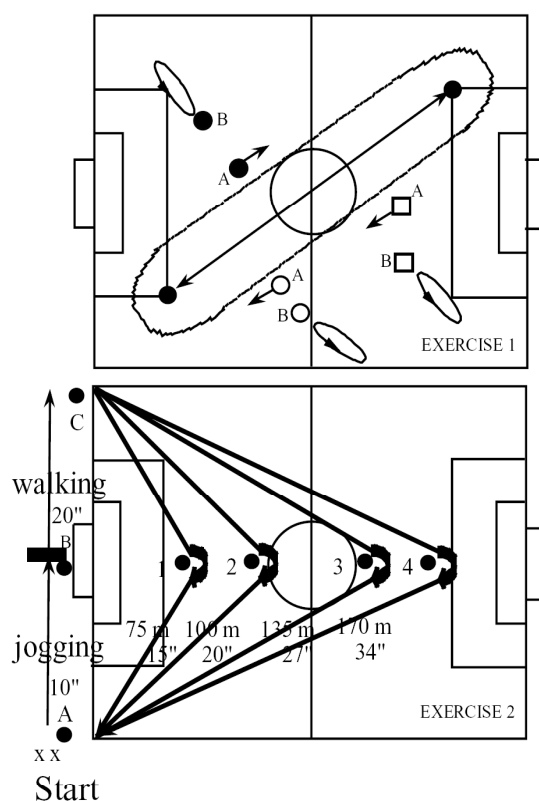


Figure 1. Illustration of two different types of high-intensity pitch-training sessions performed by the referees in study 2.

Referee A has to cover 3 clockwise high-intensity runs around the diagonal course. Once this has been completed, the referees exchange roles by handing over the vest. Referee B then completes 3 clockwise runs of the diagonal, while referee A walks (or jogs) on the outside part of the diagonal running area. Once this has been completed, again the referees switch roles by exchanging the vest, and referee A now performs 2 clockwise runs of the diagonal. The referees then repeat this sequence down to 1 clockwise diagonal run and then reverse the protocol to 1, 2, and 3 diagonal runs, respectively (resulting in 12 diagonal runs of 150 m each, or approximately 1800 m in total).

After a 2- to 3-minute recovery break, the same diagonal run is now performed but counterclockwise. Each referee now has to cover 1 round at the first run, then 2, 3 and again 3, 2, 1 (for 12 diagonal runs of 150 m each or approximately 1800 m in total). The running was performed at 90% HR_{max} , as the main training objective was high intensity (85% to 95% HR_{max}).

Exercise 2 consists of three sets of the following high-intensity runs. The referees start jogging in pairs from the start position (A) to cone B, then walk from cone B to cone C. From cone C, they then run at 90% HR_{max} around cone 1 back to the start position. For lap 2, the referees repeat the sequence from cone A to cone C, but this time they turn around cone 2, on lap 3 around cone 3, and on lap 4 around

cone 4. They then perform the same exercise but in reverse order so that set 1 consisted of 8 high-intensity runs, of various durations, at 90% HR_{max} . A 5-minute active-recovery break was taken after set 1 was completed, and then the whole procedure was repeated 3 times giving a total of 24 runs of varying distances at 90% HR_{max} over the whole training session.

Along with the two pitch-based high-intensity training sessions, two specific interval training sessions at 90% HR_{max} are also presented. These sessions were performed around a running track, and the structure of these sessions was as follows:

Exercise 3

30 seconds at 90% HR_{max} followed by 30 seconds of active recovery

60 seconds at 90% HR_{max} followed by 60 seconds of active recovery

90 seconds at 90% HR_{max} followed by 90 seconds of active recovery

120 seconds at 90% HR_{max} followed by 120 seconds of active recovery

90 seconds at 90% HR_{max} followed by 90 seconds of active recovery

60 seconds at 90% HR_{max} followed by 60 seconds of active recovery

30 seconds at 90% HR_{max} followed by 30 seconds of active recovery

Two sets, with 5-minute active recovery in between sets

Exercise 4

30 seconds at 90% HR_{max} followed by 30 seconds of active recovery

45 seconds at 90% HR_{max} followed by 45 seconds of active recovery

60 seconds at 90% HR_{max} followed by 60 seconds of active recovery

75 seconds at 90% HR_{max} followed by 75 seconds of active recovery

Repeated twice, followed by 5-minute active recovery before set 2:

75 seconds at 90% HR_{max} followed by 75 seconds of active recovery

60 seconds at 90% HR_{max} followed by 60 seconds of active recovery

45 seconds at 90% HR_{max} followed by 45 seconds of active recovery

30 seconds at 90% HR_{max} followed by 30 seconds of active recovery

This sequence was also repeated twice.

Training Monitoring

As previously mentioned, from the beginning of the 2001/2002 season, the referees received a training plan, sent via e-mail, at the start of every week. Obviously, the structure in terms of training intensity, frequency, and duration varied with the stage of the season and also with the game demands of the referees. However, in general the plans consisted of four training sessions per week. The first training session was a general recovery from the weekend match, sessions 2 and 3 were high-intensity intermittent sessions, with session 3 periodically being speed endurance. The final training session of the week, training session 4, was always a match-preparation session where the focus was on short sprints with sufficient recovery in between.

Heart rates were used as the tool to prescribe and control the training load of these sessions, as this method has been demonstrated to be a reliable index of cardiovascular strain.^{1,18}

For the international referees, adherence to the training plan was mandatory, and performance during weekly training sessions was analyzed by means of a monthly training diary filled in by the referees in which the mean and maximum heart rates during the training sessions, and also the number and distribution of their weekly training sessions, were recorded. Following analysis of the individual heart rate data, the referees were informed as to whether the aims of their training sessions had been met in terms of time spent working within the appropriate activity category. Every Tuesday during the preseason and competitive season, the international referees attended a group training session at the Katholieke Universiteit Leuven. On these occasions, the referees performed a high-intensity intermittent session, such as those presented previously, and these sessions were designed and strictly administered by the second author of this study.

The elite referees were instructed to follow the weekly training plans and also to keep a training diary. However, this group did not receive any feedback with regard to their training diaries and heart rate data obtained during their training sessions. The group training sessions for the elite referees took place every 6 weeks at the Katholieke Universiteit Leuven, where again the referees performed a high-intensity intermittent session designed and strictly administered by the second author of this study.

Yo-Yo Intermittent Recovery Test

To assess the long-term efficacy of the referees' training programs, the referees were assessed using the Yo-Yo intermittent recovery test. Specific details of this are given in Krstrup et al.¹³ The referees were assessed on four separate occasions: November 2001, January 2002, October 2002, and March 2003.

Data Reduction

During data analysis, when data appeared to be recorded outside the normal physiological range expected for the match officials, that is, more than 210 b.min⁻¹ (beats per minute) or

below 40 b.min⁻¹, the data were corrected using the error correction algorithm within the Polar Precision Performance software and checked by visual inspection of the data. If there were still periods of more than 30 seconds that remained uncorrected after this procedure, then the data were excluded from analysis.

The referees' HR_{max} was calculated using the highest 5-second peak value observed during matches, high-intensity/speed endurance training sessions, or the Yo-Yo intermittent recovery test, which has been demonstrated to produce a maximal physiological response.¹³ Using the highest 5-second peak obtained during either practice, tests, or competition enables a more accurate determination of HR_{max} than a single maximal fitness test, as Palmer et al.¹⁵ and Helsen and Bultynck⁹ reported consistently higher HR_{max} values during competition when compared to maximal laboratory tests.

For the determination of the various heart rate zones relative to the referees' HR_{max} , a distinction was made between five different activity categories that were based on findings within the football literature⁴⁻⁵ and were also in line with previous studies on the physiological demands of football refereeing.^{7,9} These categories were the following: 1) maximal effort (ME, >95% HR_{max}), 2) high intensity (HI, 86% to 95% HR_{max}), 3) low intensity (LI, 76% to 85% HR_{max}), 4) active recovery (AR, 65% to 75% HR_{max}), and 5) passive recovery (PR, <65% HR_{max}).

Data Analysis

Differences in the referees' mean and peak heart rates (expressed as a percentage of HR_{max}), and also exercise duration, between the high-intensity pitch- and track-training sessions were determined by a paired *t*-test. To further explore impact differences between two types of training sessions, a two-way analysis of variance (ANOVA) was performed using a 2 (training) (pitch sessions, track sessions) by 5 (activity category) (ME, HI, LI, AR, and PR) design with repeated measures. For the analysis of the distance covered on the Yo-Yo test, a two-way ANOVA was performed using a 2 (group) (international referees, elite referees) by 4 (testing period) (four separate testing occasions) repeated measures design. When a significant interaction was detected, the data were subsequently analyzed using a Scheffe test. Significance was set at $P < 0.05$.

RESULTS

There were no significant differences between the mean percentage HR_{max} recorded by the referees during the pitch- and track-training sessions (86.4 ± 2.9 versus 88.2 ± 2.4 % HR_{max} , $P > 0.05$) despite the pitch sessions being shorter in duration than the track sessions (19.9 ± 0.5 versus 27.0 ± 5.19 minutes, $P < 0.05$). However, there were differences between the peak heart rates recorded during the sessions, with the peaks being higher during the track sessions (94.8 ± 2.6 versus 97.4 ± 2.1 % HR_{max} , $P < 0.05$).

The analysis of the time spent working within the five different heart rate zones showed a significant effect only

TABLE 1
Mean Values and Standard Deviations of the Distance Covered on the Yo-Yo Intermittent Recovery Test for Both Belgian International ($n = 7$) and Elite Referees ($n = 12$) (in meters)

	November 2001	January 2002	October 2002	March 2003
International referees				
Mean	1720 ^a	1735 ^a	2167 ^a	2330 ^a
SD	276	318	234	268
Elite referees				
Mean	1290	1431	1777	1985
SD	407	409	393	279

^a Significant difference between international and elite referees ($P < 0.05$).

for activity category ($F_{4,104} = 144.52$, $P < 0.0001$). In line with the training objectives, the referees spent significantly more time training within the high-intensity heart rate zone than in the other four heart rate zones (Fig. 2).

Table 1 provides an overview of the means \pm SD of the distances covered by the referees during the Yo-Yo intermittent recovery test. The results of the statistical analysis showed a significant effect only for group ($F_{1,16} = 6.152$, $P < 0.0246$) and testing period ($F_{1,3} = 35.330$, $P < 0.0001$). Specifically, international referees outperformed the elite referees on each testing occasion.

From Figure 3, it becomes clear that the mean distance covered on the test increased significantly from test 1 to test 3 ($P < 0.01$), from test 2 to test 3 ($P < 0.01$), test 1 to test 4 ($P < 0.01$), and test 2 to test 4 ($P < 0.01$). However, mean distance covered did not increase between test 1 and test 2 ($P > 0.05$) and between test 3 and test 4 ($P = 0.0579$), although this value did approach conventional levels of statistical significance.

Figure 4 presents a comparison of a referee's heart rate response from test 1 and test 4. It is clear that the improvement in performance by this particular referee was, as expected, concomitant with a decrease in submaximal heart rates throughout the duration of the test along with an increase in the amount of work performed while working at HR_{max} .

DISCUSSION

In this study, we examined the long-term effect of structured training, and more specifically, high-intensity training sessions, on football referees' fitness levels. We used heart rate analysis to determine the physical load imposed on the referees during training sessions as this method has been demonstrated to be an accurate indicator of overall cardiovascular strain^{10,18} and is the most commonly used method of assessing exercise intensity used in the field.¹ To determine the impact of such training sessions, we used the Yo-Yo intermittent recovery test as this test has been

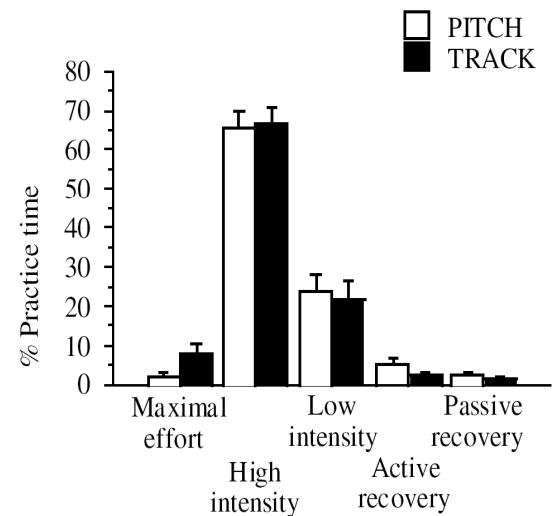


Figure 2. Means and standard errors of the percent practice time spent in the different activity categories during high-intensity training sessions.

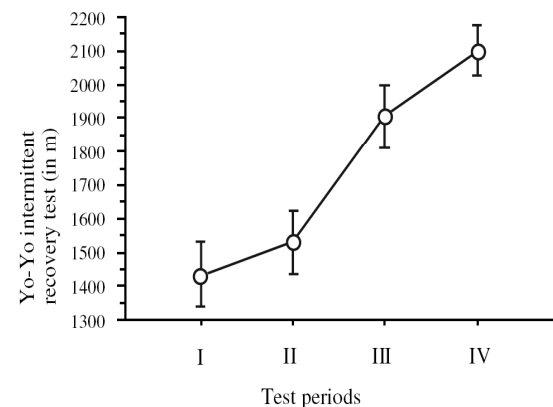


Figure 3. Mean values and standard errors of the distance covered during the Yo-Yo intermittent recovery test for the total group of match officials.

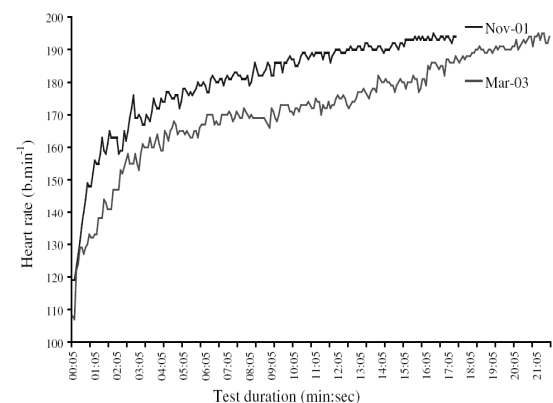


Figure 4. An example of a referee's heart rate response during two separate testing occasions on the Yo-Yo test.

reported to be highly correlated to referees' match performance.¹¹ To our knowledge, this is the first study to report actual heart rate responses during specific intensive intermittent training sessions and to also examine the long-term impact of such training sessions on referees' fitness levels.

The results obtained in the present study illustrate that both the pitch- and track-training sessions were equally successful in imposing a high-intensity training load on the referees as the mean heart rates for both types of training session fell within the high-intensity category highlighted by previous authors.^{4,5,9} Therefore, for those referees who do not have the opportunity to train together in groups or have access to pitches (geographical, work constraints, and so forth), an appropriate high-intensity training session can be achieved through specific track running. From a coaching perspective, it was also encouraging to see the lack of variability in the heart rate recordings during the training sessions from referee to referee.

There was a relatively small difference between the percentage of time spent working within the maximal effort and high-intensity activity categories during the training sessions and the data reported during the UEFA 2000 Championships⁹ (71.3% versus 60.7% of total time, respectively) and also within the scientific literature.^{7,11,12,21} Therefore, despite the fact that the heart rates during matches may represent an exaggeration in terms of the physical load placed on the referees, due to emotional stress, rising core temperature, and/or dehydration,¹¹ the similarity between these data suggests that the training sessions have a high level of cardiovascular specificity with regard to the match demands of refereeing. For the referees to achieve optimal levels of physical fitness, it is crucial that they train at and above the intensity at which they referee as training adaptations are highly specific to the type of activity and to the volume and intensity of the exercise performed.²³

The recommended guidelines for maintaining and improving aerobic fitness have been reported to be an intensity of training between 55%/65% and 90% of HR_{max} , with the lower intensity values being most applicable to individuals who are quite unfit, at a frequency of 3 to 5 times per week.² However, Krstrup and Bangsbo¹¹ and Helgerud et al.⁸ have since reported that specific high-intensity aerobic training, involving a session total of 16-minute high-intensity running performed 2 to 3 times per week, improved fitness levels in football referees and players, respectively. Specifically, for referees, Krstrup and Bangsbo¹¹ reported that following a 12-week period of intense, intermittent training (90% HR_{max}) performed 3 times per week, the distance covered by referees on the Yo-Yo test improved by 31%. For players, Helgerud et al.⁸ reported that 8 weeks of high-intensity training sessions, performed twice per week at 90% to 95% HR_{max} , brought about a significant increase in VO_{2max} , blood lactate threshold, and running economy in football players.

With this in mind, we have demonstrated that the high-intensity pitch- and track-training sessions presented in this study are of a suitable intensity (86.4 ± 2.9 and 88.2 ± 2.4 % HR_{max} , respectively) and duration (19.9 ± 0.5 and 27.0

± 5.19 minutes, respectively) to bring about significant training adaptations in referees. However, it is important to highlight that when the training load becomes too intense, the body's ability to recover is impaired, which may lead to underperformance.²³ Therefore, in line with Krstrup and Bangsbo¹¹ and Helgerud et al.,⁸ we recommend that the referees do not perform more than two to three high-intensity training sessions within 1 week.

Krstrup et al.¹² reported that games are characterized by long, low-intensity activity periods interspersed with frequent brief intense exercise periods. Therefore, speed endurance and good recovery in between intense periods of exercise are vital for referees and assistant referees.¹² Consequently, the training programs that referees follow on a weekly basis should have a blend of high- and low-intensity aerobic sessions, complemented with training sessions dedicated toward the improvement of speed and the ability to perform repeated bouts of intense exercise with incomplete recovery (speed endurance).

Previous work¹¹ has examined only the short-term (12 weeks) effect of intense, intermittent training in relation to referees' fitness levels. However, our results clearly demonstrate that over a period of almost two football seasons, following a structured training regime that replicated match demands through intermittent high-intensity training, referees significantly improved their aerobic and anaerobic fitness levels as both energy systems have been reported to be highly taxed during the Yo-Yo intermittent recovery test.¹³ What is even more encouraging was that both groups of referees continued to progressively improve their fitness levels with each test.

The international referees had significantly higher levels of fitness at the start of the study when compared to the elite referees. This finding is consistent with Rontoyannis et al.¹⁹ who also reported that fitness levels were lower among those referees from a lower standard of competition. It has been demonstrated that the initial level of fitness largely governs the training response with relative improvements being greater for less fit individuals.²³ Our results are consistent with this finding as the elite group of referees, whose initial level of fitness was significantly lower than the international referees, demonstrated an average 53.9% improvement from test 1 to test 4. However, as a result of comprehensive feedback of their training diaries and training data, along with weekly group-training sessions, the international referees still improved their performance on the Yo-Yo intermittent recovery test by an average 35.5% from test 1 to test 4.

Despite demonstrating a higher initial fitness level when compared to the elite referees, the international group of referees also managed to improve their performance on the Yo-Yo test with every subsequent testing session. Also, the distances attained on the Yo-Yo intermittent recovery test during tests 3 and 4 were highly comparable with professional football players as Krstrup et al.¹³ reported a mean \pm SD distance covered during the Yo-Yo intermittent recovery test of 2075 ± 156 m for 17 players, with a mean age of 28 years (range, 25 to 36).

However, when training referees it is important not to lose sight of the main goal, which is improving match per-

formance. It is no good improving fitness levels if it does not improve match performance. It has previously been demonstrated that intense intermittent training does not only improve fitness levels in football referees, but this improvement in fitness is concomitant with an improved match performance.¹¹ Specifically, the amount of high-intensity running performed by the referees increased by $23 \pm 8\%$, mostly because of a marked increase in high-intensity running during the second half.¹¹ Also, the number of high-intensity running bouts was higher after than before the intermittent training, and the highest recorded values of distances from match incidents were less after the intermittent training.¹¹

With this in mind, we are confident that improvements in the referees' fitness in the present study will have had a positive impact on the field of play as the Yo-Yo intermittent recovery test is highly correlated with a referees' match performance.¹¹ This improvement in match performance will have been achieved through a marked increase in the referees' ability to perform repeated intense work, thus enabling the referee to keep up with play more efficiently and obtain optimal viewing positions, especially during the second half of matches when literature has consistently demonstrated a decreased work rate by the referee during the second half of matches owing to fatigue.^{6,7,11,12,16}

The training sessions and results presented within this study may also have implications for referees involved in other field sports such as hockey and rugby where inadequate fitness levels may impair the referees' ability to keep up with play and subsequently make the correct decisions. For example, Martin et al.¹⁴ reported that rugby referees are required to have well-developed anaerobic energy systems to enable them to respond to fast passages of play and to be "on the spot" to make decisions. Therefore, the intensive intermittent training sessions performed by the football referees in the present study may help to develop appropriate levels of aerobic and anaerobic fitness levels in referees involved in other team sports.

The absence of any significant improvements between test 1 and test 2 can be attributed to the timing of the testing sessions; only 2 months separated the two tests. The fact that the testing sessions were not performed at evenly spaced intervals over the duration of the study period was a limitation associated with the study design. This was due to the problems associated with getting all the referees to the testing venue at the same time as the referees within this study have full-time professions outside that of football, and very often work and match commitments interfere with training and testing schedules. Also, another limitation was the lack of a control group against which to gauge the improvements in fitness. However, as fitness has been reported to decrease with age,^{20,23} the fact that the referees improved their performance on the Yo-Yo test from November 2001 to March 2003 by an average 46.5% clearly demonstrates the effectiveness of such training session regimes for referees. Also, the involvement of age-matched sedentary subjects as a control group would not have been representative of a football refereeing population and has not been used in previous work.¹¹

CONCLUSION

The data of the present study demonstrate that a structured training program with intensive intermittent training sessions can significantly improve the fitness levels of football referees, as determined by the Yo-Yo intermittent recovery test. The improvement in fitness levels will have helped improve the referees' match performance through an increased ability to perform high-intensity running, thus enabling them to better keep up with play and to obtain optimal viewing positions in relation to critical match incidents.

ACKNOWLEDGMENT

We would like to thank FIFA's F-MARC (Medical Assessment and Research Centre) for their financial support in this project. We also sincerely thank all the members of the FIFA and the Belgian referees' committee for their collaboration in this project. The cooperation of the Belgian match officials was of invaluable importance. We are also very grateful to Roel De Clerck for assistance with the data reduction. Finally, we want to acknowledge Dr. Jiri Dvorak and Dr. Astrid Junge for providing us with valuable comments on earlier drafts of the paper.

REFERENCES

1. Achten J, Jeukendrup A: Heart rate monitoring: Applications and limitations. *Sports Med* 33(7): 517-538, 2003
2. American College of Sports Medicine: The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults: Position stand. *Med Sci Sports Exerc* 30(6): 1998
3. Bangsbo J, Norregaard L, Thorsoe F: Activity profile of competition soccer. *Int J Sports Med* 13: 125-132, 1991
4. Bangsbo J: The physiology of soccer—with special reference to intense intermittent exercise. *Acta Physiol Scand* 151(suppl): 619, 1994
5. Bangsbo J: Energy demands in competitive soccer. *J Sports Sci* 12: S5-S12, 1994
6. Catterall C, Reilly T, Atkinson G, et al: Analysis of the work rates and heart rates of association football referees. *Br J Sports Med* 27: 193-196, 1993
7. D'Ottavio S, Castagna C: Analysis of match activities in elite soccer referees during actual match play. *J Sports Med Phys Fitness* 41: 27-32, 2001
8. Helgerud J, Engen LC, Wisloff U, et al: Aerobic endurance training improves soccer performance. *Med Sci Sports Exerc* 33(11): 1925-1931, 2001
9. Helsen WF, Bultynck JB: Physical and perceptual-cognitive demands of top-class refereeing in association football. *J Sports Sci*: In press
10. Johnston L, McNaughton L: The physiological requirements of soccer refereeing. *Aust J Sci Med Sport* 26(3/4): 67-72, 1994
11. Krstrup P, Bangsbo J: Physiological demands of top-class soccer refereeing in relation to physical capacity: Effect of intense intermittent exercise training. *J Sports Sci* 19(11): 881-891, 2001
12. Krstrup P, Mohr M, Bangsbo J: Activity profile and physiological demands of top-class soccer assistant refereeing in relation to training status. *J Sports Sci* 20: 861-871, 2002
13. Krstrup P, Mohr M, Amstrup T, et al: The Yo-Yo intermittent recovery test: Physiological response, reliability and validity. *Med Sci Sports Exerc* 35(4): 697-705, 2003

14. Martin J, Smith N, Tolfrey K, et al: Activity analysis of English premier-ship rugby football union refereeing. *Ergonomics* 44(12): 1069–1075, 2001
15. Palmer G, Hawley JA, Dennis SC, et al: Heart rate responses during a 4-d cycle stage race. *Med Sci Sports Exerc* 26(10): 1278–1283, 1994
16. Rebelo N, Krstrup P, Soares J, et al: Reduction in intense intermittent exercise performance during a soccer match. *J Sports Sci* 16: 482–483, 1998
17. Reilly T, Thomas V: A motion analysis of work-rate in different positional roles in professional football match play. *J Hum Mov Stud* 2: 87–97, 1976
18. Reilly T: Energetics of high intensity exercise (soccer) with particular reference to fatigue. *J Sports Sci* 15: 257–263, 1997
19. Rontoyannis GP, Stalikas A, Sarros G, et al: Medical, morphological and functional aspects of Greek football referees. *J Sports Med Phys Fitness* 38(3): 208–214, 1998
20. Stamford BA: Exercise and the elderly. *Exerc Sport Sci Rev* 16: 341–379, 1988
21. Weston M, Brewer J: An investigation into the physiological demands of soccer refereeing. *J Sports Sci* 20: 59–60, 2002
22. Williams AM, Lee D, Reilly T: *A Quantitative Analysis of Matches Played in the 1991–92 and 1997–98 Seasons*. London, Football Association, 1999
23. Wilmore JH, Costill DL: *Physiology of Sport and Exercise*. Second edition. Champaign, IL, Human Kinetics, 1998

CHAPTER 3.0

MATCH DEMANDS

3.1

STUDY II

Weston, M., Bird, S., Helsen, W., Nevill, A. & Castagna, C. (2006) The effect of match standard and referee experience upon the objective and subjective match workload of English Premier League referees. *Journal of Science and Medicine in Sport*, 9, 256-262.



ELSEVIER

ORIGINAL PAPER

The effect of match standard and referee experience on the objective and subjective match workload of English Premier League referees

M. Weston^a, S. Bird^b, W. Helsen^{a,*}, A. Nevill^c, C. Castagna^d

^a Department of Biomedical Kinesiology, Katholieke Universiteit Leuven, Tervuursevest 101, 3001 Heverlee (Leuven), Belgium

^b Centre for Population Health in the West, Sunshine Hospital, St. Albans, Vic., Australia

^c School of Sport, Performing Arts and Leisure, University of Wolverhampton, Walsall, United Kingdom

^d School of Sport and Exercise Sciences, Faculty of Medicine and Surgery, University of Rome Tor Vergata, Italy

Accepted 22 March 2006

KEYWORDS

Soccer;
Refereeing;
Heart rates;
Ratings of perceived exertion;
Competition

Summary The aim of the present study was to examine the effect of match standard and referee experience on the objective and subjective workload of referees during English Premier League and Football League soccer matches. We also examined the relationship between heart rate (HR) and ratings of perceived exertion (RPE) for assessing match intensity in soccer referees. Heart rate responses were recorded using short-range telemetry and RPE scores were collected using a 10-point scale. Analysis revealed a significant relationship between mean match HR and match RPE scores ($r=0.485$, $p<0.05$, $n=18$). There were significant differences in match HR (Premier League $83.6 \pm 2.6\%$ maximal HR (HR_{max}) versus Football League $81.5 \pm 2.2\%HR_{max}$, $p<0.05$) and match RPE scores (Premier League 7.8 ± 0.8 versus Football League 6.9 ± 0.8 , $p<0.05$) between standards of competition. Referee experience had no effect on match HR and RPE responses to Premier League and Football League matches. The results of the present study demonstrate the validity of using HR and RPE as a measure of global match intensity in soccer referees. Referee experience had no effect on the referees' objective and subjective match workload assessments, whereas match intensity was correlated to competition standard. These findings have implications for fitness preparation and evaluation in soccer referees. When progressing to a higher level of competition, referees should ensure that appropriate levels of fitness are developed in order to enable them to cope with an increase in physical match demands.

© 2006 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.

* Corresponding author. Tel.: +32 16 32 90 68; fax: +32 16 32 91 97.
E-mail address: werner.helsen@faber.kuleuven.be (W. Helsen).

Introduction

Because of the practicality of monitoring referees' heart rates during the course of a soccer match as they do not experience physical contact,¹ heart rates (HR) responses during competitive matches have already been reported in the literature.^{2–4} Krstrup and Bangsbo⁴ reported that, over the duration of selected Danish League matches, the referees' HR was calculated to be $85 \pm 1\%$ of maximal HR (HR_{max}). This figure was consistent with the $85 \pm 5\%HR_{max}$ reported for top-class referees during the 2000 European Soccer Championships.³ However, Krstrup and Bangsbo⁴ reported that HR recordings obtained on referees during a match may overestimate the physical intensity of matches because such factors as isometric contractions, thermal and emotional stress can elevate HR beyond the normal HR- VO_2 relationship. Indeed, Jeukendrup and van Diemen⁵ reported that HR recordings during competitive exercise might be a better indicator of whole body metabolic stress as opposed to exercise intensity.

Ratings of perceived exertion (RPE) have been widely used to assess exercise intensity and when the scale is used correctly it is a very accurate method for monitoring exercise intensity.⁶ Furthermore, Foster⁷ reported that the RPE technique is consistent with objective physiological indices of exercise training. However, the relationship between subjective data obtained using this technique and objective data, acquired using HR recordings, has yet to be examined with regard to the assessment of global match intensity in elite soccer referees.

Authors have reported that soccer match intensity is correlated with the level of competition.^{8,9} Mohr et al.⁹ reported that top-class players performed more high-intensity running during matches, a valid measure of physical match performance in soccer, than moderate professional players did. This difference in intensity with match standard may impact on referees' objective and subjective match workload responses when refereeing at different levels of competition. However, the effect of competition standard on match intensity in elite soccer referees has yet to be thoroughly examined. Catterall et al.,² using data collected from Premier League, Football League and non-Football League matches, reported that the mean HR was unaffected by the category of competition, although higher peak HRs were recorded during games in the top division. Unfortunately, HR data were only collected from one match per referee. Therefore, as the referees did not officiate at more than one standard of competition, an accu-

rate assessment of the effect of match standard on match intensity remains relatively unexplored.

Wilkins et al.¹⁰ observed high HRs during periods of low physiological stress when psychological stress was likely to be high, such as restraining players and intently watching play. As mental stress can elevate HR through the release of such hormones as adrenaline and noradrenaline,¹¹ referee experience may also have an impact on the HR responses during matches as psychological stress may be reduced as officials gain more experience.¹⁰ Krstrup and Bangsbo⁴ reported few significant differences in observed match activities between top-class (international) and high-standard (national) Danish referees. However, the referees' HRs were measured on only two occasions. To our knowledge, no attempt has been made within the literature to examine the effect of experience on referees' match HRs and also RPE scores.

Therefore, the objectives of the present study were: (i) to examine the relationship between HR and RPE during soccer matches and (ii) to investigate the effect of match standard and referee experience on referees' subjective and objective ratings of global match intensity.

Method

Participants

In English soccer, the highest level of competition is the Football Association (FA) Premier League. The Football League is the next level down and comprises three separate divisions with a progressive decrease in standard from the First to the Third Division. Data were collected from 18 of the 20 full-time, professional English FA Premier League referees who were comfortable wearing a HR monitor during their matches. The match HR and RPE data were collected as part of the ongoing sports science support provided to the Premier League referees by the first author of this study, whereby objective and subjective ratings of exercise intensity are used to evaluate the referees' overall training and match loads. The mean age and experience, in terms of years refereed in the Premier League, was 41.8 year (range 33–47 year) and 5.4 year (range 2–11 year), respectively.

Assessment of RPE

Each referee was asked to assess and record the RPE following each of their matches during the 2002–2003 English football season, giving a total of 527 match observations (median of 31 matches,

range 6–41). The referees recorded their RPE score 30 min after the match had ended in order to obtain a global intensity rating for the entire match. Therefore, a particularly easy or difficult bout of exercise towards the end of the match would not dominate the referees' rating, as per Foster⁷ and Foster et al.¹² The scale used in the present study was the category ratio scale (CR10-scale) from Borg et al.⁶ and has been reported to correlate significantly with objective indices of exercise training such as HR and blood lactate.¹² The referees were familiar with the use of RPE scores as they had used the scale as part of their training monitoring procedures over a 12-month period prior to the study.

Heart rate sampling

The referees' HR during matches was recorded via short-range telemetry using Polar S610 watches (Polar, Kempele, Finland), with the data being recorded every 5 s. The referees were already familiar with using the HR monitor and its functions, and were also aware of how to fit the transmitter correctly. They were instructed to insert markers into the HR files, using the appropriate button on the HR monitor, in order to indicate the precise moment at which the match halves started and ended. Data recording commenced from the start of the referees' warm up and finished at the final whistle. Following each match, the data were downloaded on to a computer for analysis using the Polar Precision software, version 3.0 (Polar, Kempele, Finland). Fig. 1 provides a typical example of a referee's HR response to a Premier League match.

Study design

The referees in the present study officiate in the FA Premier League, with few selected matches during the season in the Football League. Because of the limited number of matches refereed at the lower standard of competition, HR recordings and

RPE scores from a total of six matches per referee were analysed during the 2002–2003 season. These matches consisted of three Premier League and three Football League matches in an attempt to examine the effect of match standard on the referees' HR and RPE responses during matches. The six matches were grouped into three pairs, with data collected from each Football League match taking place within 14 days of data collected from a Premier League match in an attempt to eliminate any potential impact of fitness variations on the referees' match data. To gain an accurate representation of the whole season, and to eliminate any possible effect of seasonal variation on HRs, as thermal stress can elevate HR beyond the normal HR-VO₂ relationship,⁴ the season was divided into three equal segments: August–October, November–January and February–April. Each pair of match observations was taken from each segment of the season. Also, the HR responses of eight randomly chosen referees during four matches (two Premier League and two Football League) were studied in order to determine the intra-referee variation.

The referees' experience of refereeing in the Premier League was recorded at the beginning of the 2002–2003 season. To examine the effect of referee experience on match HR and RPE responses during Premier League and Football League matches, the referees were divided into two groups: a low experience group, with four seasons or fewer refereeing in the Premier League ($n = 10$, 3.1 ± 0.9 year experience) and a high experience group with more than four seasons refereeing in the Premier League ($n = 8$, 8.4 ± 2.2 year experience).

Laboratory assessment for maximal heart rate

All of the 18 referees who participated in this study visited the laboratory in September 2002 and completed a maximal, incremental test for the determination of HR_{max} and VO_{2max}, as part of the ongoing sports science support provided to the referees. Following 5 min of self-paced warm up running, the referees performed an interval-based protocol, comprising five 4 min stages, on a motorised treadmill (Woodway, Germany). These stages were interspersed with a 30 s recovery period during which the referees stopped running and straddled the treadmill belt and the test administrator increased the running speed. The initial running speeds were selected from the perceived fitness of the referees, with each referee starting the protocol at either 6.5 or 7.1 miles h⁻¹, with a treadmill inclination of

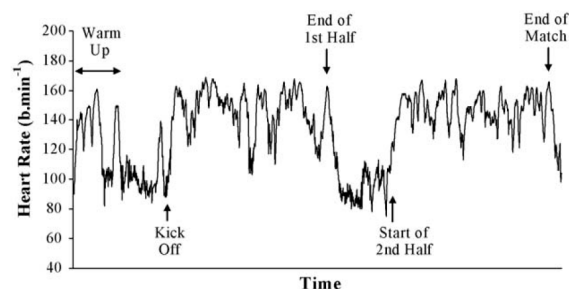


Figure 1 An illustration of a referee's HR response to a typical Premier League match.

1%. The speed was then subsequently increased by 0.6 miles h^{-1} for each of the following four stages. At the end of the fifth stage, the referees continued to run and the inclination on the treadmill was increased by 1% every 30 s until volitional exhaustion. HR was monitored via short-range telemetry using Polar S610 watches (Polar, Kempele, Finland) and the referees' HR_{max} was calculated as the highest 5 s peak value attained during the assessment.

Statistical analysis

The following dependent measures were identified and calculated:

- The mean match HR was obtained from the recordings during the entire match, including periods of time added on for stoppages at the end of each half. The value was expressed both absolute and relative to the referees' HR_{max} .
- Within-match HR: matches were divided into six periods over the duration of the match in accordance with previous studies,^{1,3} with the mean absolute and relative HRs recorded for each of these periods. These periods were from the kick-off to 15 min, 16–30 min and 31 min until the end of each 45 min half. This procedure enabled the detection of whether or not the referees modified their activity pattern during the game.¹

Pearson's product moment correlations were calculated on the relationship between RPE and the mean match HR for each individual referee. A *z*-statistic was used to test for significance of the correlation coefficient. The coefficient of variation was used as a measure of intra-referee match HR responses.¹³

A Student's paired *t*-test was used for the analysis of the Premier League referees' mean match HRs, peak match HR and match RPE scores in both Premier League and Football League matches; a repeated measures design over two occasions. To examine further the effects of match standard on match HR responses throughout the six 15 min periods, a two-way ANOVA was carried out on the HR, expressed as $\%\text{HR}_{\text{max}}$ values, using a 2-match standard (Premier League, Football League) by 6-period (Periods 1–6) design with repeated measures. When a significant interaction was detected, the data were subsequently analysed using a Tukey–Kramer post hoc test. Differences between high-experience and low-experience referees' mean match HRs and match RPE scores were determined by unpaired *t*-tests. Significance was set at $p < 0.05$.

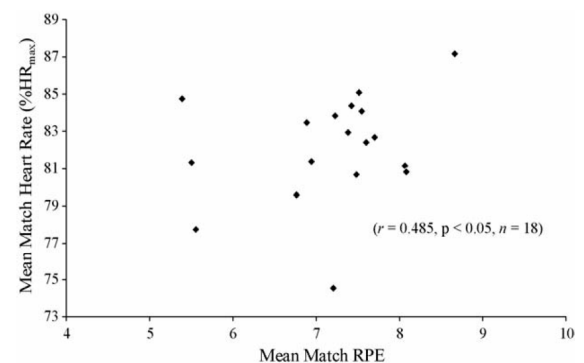


Figure 2 The relationship between mean match HRs and match RPE scores.

Results

Relationship between RPE and mean match heart rate

Fig. 2 illustrates the relationship between mean match HRs and mean match RPE scores for each referee ($n = 18$). There was a significant correlation between mean match HR and mean match RPE score ($r = 0.485$, $p < 0.05$, $n = 18$).

Match standard

The mean match HR response for the Premier League referees was higher when refereeing Premier League matches compared to Football League matches ($83.6 \pm 2.6\%\text{HR}_{\text{max}}$ versus $81.5 \pm 2.2\%\text{HR}_{\text{max}}$, $p = 0.0027$). However, there were no differences in the Premier League referees' peak HRs in Premier League and Football League matches ($95.6 \pm 2.5\%\text{HR}_{\text{max}}$ versus $95.0 \pm 2.1\%\text{HR}_{\text{max}}$, $p > 0.05$). The mean match RPE scores for the Premier League referees were also significantly higher for Premier League matches compared to Football League matches (7.8 ± 0.8 versus 6.9 ± 0.8 , $p < 0.0001$).

As can be seen from Fig. 3, the results of the 2 (match standard) \times 6 (match period) ANOVA showed a significant effect for match standard ($F_{1,34} = 6.22$, $p = 0.017$), with the within-match HRs on Premier League being higher throughout the entire match when compared to the Football League matches. There was also a significant effect for match period on the within-match HRs ($F_{5,85} = 18.90$, $p < 0.001$). Results of the post hoc analysis showed that, during both Premier League and Football League matches, HRs in the second half of matches increased from periods 4 to 6 but not for the corresponding periods during in the first half (periods 1–3). The lowest match period HR for both standards of competition

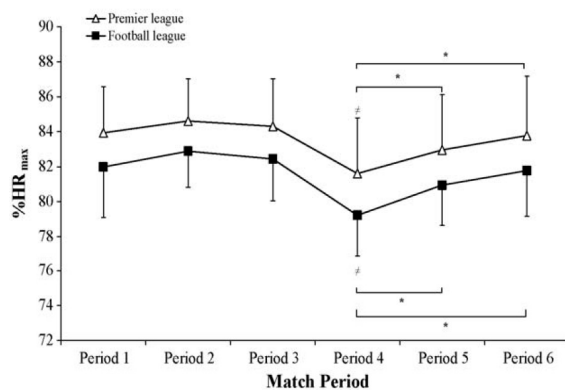


Figure 3 Referees' mean (\pm S.D.) within-match HR responses per match period during three Premier League and Football League matches; * $p < 0.001$; # statistically different from all other match periods at each standard of competition ($p < 0.05$).

was period 4, which was the first 15 min of the second half. The pattern of the HR response over the six match periods was unaffected by the standard of competition ($F_{5,170} = 0.23$, $p = 0.948$) in that the within-match HRs displayed an almost parallel trend over the duration of the Premier League and Football League matches.

Referee experience

Mean HRs for both low-experience referees ($82.7 \pm 8.6\%HR_{max}$ versus $80.8 \pm 5.6\%HR_{max}$, $p < 0.001$) and high-experience referees ($84.8 \pm 2.4\%HR_{max}$ versus $82.3 \pm 2.3\%HR_{max}$, $p = 0.016$) were higher during Premier League matches than during Football League matches. However, both Premier League ($p = 0.084$) and Football League ($p = 0.134$) mean match HRs were unaffected by referee experience. Referee experience had no effect on match RPE scores during both Premier League (7.8 ± 0.5 versus 7.7 ± 1.0 , $p = 0.872$) and Football League matches (6.8 ± 0.7 versus 6.9 ± 0.9 , $p = 0.755$).

Intra-referee variation

The HR responses of eight referees during four matches (two Premier League and two Football League) were studied in order to determine the intra-referee variation. For the Premier League matches, the mean intra-referee variation was calculated to be $2.5 \text{ beats min}^{-1}$ ($0.4\text{--}5.8 \text{ beats min}^{-1}$), with a coefficient of variation of 1.7%. For the Football League matches, the mean intra-referee variation was observed to be $2.4 \text{ beats min}^{-1}$ ($0.4\text{--}6.0 \text{ beats min}^{-1}$), with a coefficient of variation of 1.7%.

Discussion

In this study, we assessed the validity of using HRs as an indicator of global match intensity in soccer referees by examining the relationship between match HRs and match RPE scores. Ratings of perceived exertion have been reported to be a valid, subjective estimate of physical load during non-steady state exercise, including very high-intensity interval training and team sport practice and competition.^{7,12} In the present study, mean match HRs were significantly correlated with RPE scores. Therefore, despite HRs having previously been reported to overestimate slightly the physical load placed on referees during matches,⁴ our results demonstrate that match HR may be considered a valid indicator of global match intensity in this group of elite-level soccer referees. Also, a low coefficient of variation in match HRs suggests that HR can be regarded as a valuable parameter to evaluate intra-referee match intensity. Because of the practicality and cost effectiveness of such intensity assessment techniques, they can be used by referees at all levels of the game in order to assess the subjective and objective workload imposed on them during matches.

The fact that the referees' HRs increased significantly from the first to the last period of the second half for both standards of competition may be attributed to cardiovascular drift. Krstrup and Bangsbo⁴ reported that dehydration accounted for a 5 beats min^{-1} rise in referee HR over the duration of a half. Also, Reilly¹⁴ reported that the distribution of goals shows a bias towards more goals being scored at the end of the game. This final burst of activity by the players may have, in turn, led to an increase in activity by the referees which would be manifested through an increased HR.

During both Premier League and Football League matches, the first 15 min period of the second half produced the lowest HR recordings. Literature has demonstrated, that for both soccer players^{9,15} and referees,^{1,4} match intensity is lower during the second half of matches, especially during the initial stages of the second half. Also, Krstrup et al.¹⁶ reported that the quadriceps muscle temperature of match officials was lowered by more than 1°C following the half-time interval and that this was concomitant with a reduced physical performance during the initial period of the second half.

Premier League and Football League match HRs and RPE scores were not related to refereeing experience. However, match HRs and RPE scores recorded from the referees were significantly higher during Premier League matches compared to the Football League. As match HRs have

been significantly correlated with subjective ratings of exercise intensity, it appears that, for this group of elite-level referees, global match intensity is correlated to the standard of competition refereed, as both subjective and objective measures of physical workload were significantly higher during the higher standard of competition. This finding is consistent with the research on players, which has demonstrated that soccer match intensity is correlated with the level of competition, with higher standard matches being more intensive.^{8,9}

As global match intensity was correlated to the level of competition, this clearly has an implication for a referee's fitness preparation and evaluation. Referees and football governing bodies should ensure that appropriate levels of match fitness are developed in order to enable referees to cope with the increase in match intensity that accompanies an elevation in competition standard.

The measurement, and subsequent analysis, of HR and RPE can provide referees and training experts with useful implications for training monitoring, thus ensuring that referees can cope with the physical demands of their matches at all levels of competition. HRs have been used to detect overtraining at an early phase, either at rest or during a maximal assessment, and also during sleep.⁵ Session RPE scores, defined as session duration multiplied by the RPE score, enable accessory indices of training to be calculated, such as monotony and strain.⁷ Therefore, the combination of HR and RPE monitoring during training and competition may allow a referee to achieve the goals of training while minimising undesired training outcomes.

In conclusion, the relationship between match HRs and RPE indicates that HRs may be considered as a valid measure of referees' workload during competitive matches. Experience had no effect on the referees' objective and subjective match workload assessment. However, the referees' global match intensity recordings were correlated to the level of competition as the referees recorded higher HRs and RPE scores for the higher standard of competition. This has clear implications for the fitness preparation and evaluation of referees when progressing to a higher standard of competition.

Practical implications

- Match intensity in Premier League referees is correlated with the level of competition, with higher standard matches being more intensive.

- Referees and football governing bodies should ensure that appropriate levels of match fitness are developed in order to enable referees to cope with the increase in match intensity that accompanies an elevation in competition standard.
- Despite an increase in global match intensity at the higher standard of competition, objective and subjective ratings of exercise intensity during matches are unaffected by referee experience.

Acknowledgements

We would like to thank FIFA's F-MARC (Medical Assessment and Research Centre) for its financial support in this project. We also sincerely thank Philip Don and Keith Hackett (FA Premier League Referees Officers) for their help and support. The cooperation of the Premier League referees was of invaluable importance.

References

1. D'Ottavio S, Castagna C. Analysis of match activities in elite soccer referees during actual match play. *J Sports Med Phys Fitness* 2001;15(2):167–71.
2. Catterall C, Reilly T, Atkinson G, et al. A. Analysis of the work rates and heart rates of association football referees. *Br J Sports Med* 1993;27:193–6.
3. Helsen WF, Bultynck JB. Physical and perceptual-cognitive demands of top-class refereeing in association football. *J Sports Sci* 2004;22(2):179–91.
4. Krstrup P, Bangsbo J. Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. *J Sports Sci* 2001;19(11):881–91.
5. Jeukendrup A, van Diemen A. Heart rate monitoring during training and competition in cyclists. *J Sports Sci* 1998;16:S91–9.
6. Borg G, Hassmen P, Lagerstrom M. Perceived exertion related to heart rate and blood lactate during arm and leg exercise. *Eur J Appl Physiol Occup Physiol* 1987;56:679–85.
7. Foster C. Monitoring training in athletes with special reference to overtraining syndrome. *Med Sci Sports Exerc* 1998;30(7):1164–8.
8. Ekblom B. Applied physiology of soccer. *Sports Med* 1986;3:50–60.
9. Mohr M, Krstrup P, Bangsbo J. Match performance of high-standard soccer players with special reference to development of fatigue. *J Sports Sci* 2003;21:519–28.
10. Wilkins HA, Petersen SR, Quinney HA. Time-motion analysis of and heart rate responses to amateur ice hockey officiating. *Can J Sports Sci* 1999;16(4):302–7.
11. Astrand PO, Rodahl K. *Textbook for work physiology*. New York, USA: McGraw-Hill; 1977.
12. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. *J Strength Cond Res* 2001;15(1):109–15.

13. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;**26**:217–38.
14. Reilly T. Motion analysis and physiological demands. In: *Science and soccer*. London: E & FN Spon.; 1996. pp. 65–81.
15. Bangsbo J, Norregaard L, Thorsoe F. Activity profile of competition soccer. *Can J Sports Sci* 1991;**16**:110–6.
16. Krstrup P, Mohr M, Bangsbo J. Activity profile and physiological demands of top-class soccer assistant refereeing in relation to training status. *J Sports Sci* 2002;**20**:861–71.

Available online at www.sciencedirect.com



3.2

STUDY III

Weston, M., Castagna, C., Impellizzeri, F., Rampinini, E. & Abt, G. (2007)
Analysis of physical match performance in English Premier League soccer
referees with particular reference to first half and player work rates. *Journal
of Science and Medicine in Sport*, 10 (6), 390-397.



ELSEVIER

ORIGINAL PAPER

Journal of
Science and
Medicine in
Sport

www.elsevier.com/locate/jsams

Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates

Matthew Weston^{a,*}, Carlo Castagna^b, Franco M. Impellizzeri^c,
Ermanno Rampinini^c, Grant Abt^d

^a Professional Game Match Officials Limited, The FA Premier League, 30 Gloucester Place, London, UK

^b School of Sport and Exercise Sciences, Faculty of Medicine and Surgery,
University of Roma Tor Vergata, Rome, Italy

^c Human Performance Laboratory, S.S.MAPEI srl, Via Don Minzoni 34, Castellanza, Varese, Italy

^d St. Martin's College, Lancaster, UK

Received 12 May 2006; received in revised form 14 September 2006; accepted 21 September 2006

KEYWORDS

Association football;
Match analysis;
Fatigue;
Intermittent exercise

Summary The aim of the present study was to examine the influence of first half activity, overall match intensity and seasonal variation on the physical match performances of English Premier League football referees. Match analysis data was collected using the Prozone[®] match analysis system from 19 fulltime professional referees during a total of 254 matches in the 2004–2005 season. Physical match performances were classified into three separate categories: 1, total distance covered (TD); 2, high-intensity running distance (running speed > 5.5 m/s, HIR); 3, average distance from infringements (DI). Using these match activity variables the influence of first half TD and HIR distances on second half activities and also the influence of players' match activities upon the referees' physical match performances were examined. The main finding of the present study was that the physical match performances of the referees were partly related to those of the players, in that the referees' HIR correlated with players' HIR ($r=0.43$, $p<0.0001$, $n=212$). Furthermore, first half TD and HIR distances were found to be related to second half coverage in referees ($r=0.47$ and $r=0.52$, respectively, $p<0.001$, $n=254$). These results demonstrate a need to assess the overall match intensity prior to examining the physical match performance of the referee. Further examination is required as to whether reduced physical performances in the second half of matches are a consequence of referee fatigue, tactical strategies on behalf of the referee or reduced player match activities resulting in a slower tempo of match.

© 2006 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.

* Corresponding author. Tel.: +44 2078649128.

E-mail address: mweston@fapl.co.uk (M. Weston).

Introduction

The physical demands imposed upon football referees during matches have been well reported within the scientific literature over recent years. Catterall et al. evaluated the work rate of 14 English football referees and reported a mean distance covered over the period of a whole match of 9438 m, with a range of 7977–10,187 m.¹ More recently, D'Ottavio and Castagna reported mean match distance coverage of $11,469 \pm 983$ m for experienced referees officiating in the top Italian football leagues.²

In terms of match demands and the subsequent development of fatigue during matches, the TD covered by a football referee has been reported to be a poor measure of the physical stress imposed by a match since standing, walking and jogging account for more than 75% of total match activity,³ whereas the amount of HIR has been reported to be a better indicator of the physical demands of a match and in turn the development of fatigue.⁴ Krstrup and Bangsbo reported intra-individual coefficients of variation for match distance and HIR of 4 and 6%, respectively.³ However, the factors that may impact upon a referee's physical match performance remain relatively unexamined.

Stølen et al., in a review article on the physiology of football, reported that the analysis of between-half distance coverage is of great interest as it can reveal the occurrence of fatigue and/or refereeing strategies; both of which would be classified as determinants of overall physical match performance.⁵ However, the literature presents conflicting evidence with regard to this aspect of physical match performance in referees. Catterall et al. reported that less TD was covered by the referees in the second half compared to the first half¹ and D'Ottavio and Castagna reported a 4% decrease in TD covered in top level Italian referees.² However, in contrast to these findings, Krstrup and Bangsbo reported no significant difference in TD coverage between the two halves.³

As previously reported, the amount of HIR during a match may reveal more relevant information in the assessment of fatigue development during a game.⁵ Johnston and McNaughton⁶ and also Krstrup and Bangsbo³ have reported that the amount of HIR (running speed $> 15.0 \text{ km h}^{-1}$) performed by football referees decreased during the second half of a match. D'Ottavio and Castagna, however, observed no between-half differences in HIR (running speed $> 18.1 \text{ km h}^{-1}$) despite a decrease in TD covered.² Castagna and Abt⁷ supported this evidence of possible 'sparing behaviour' on behalf of the referee. Therefore, it is difficult to make any unequivocal conclusions with regard to the impact

of first half performance on second half performance in football referees.

The intensity at which a match is played could well influence the work rate of the referee. Catterall et al. stated that the overall pattern of a referee's match activity was acyclical but varied in parallel with the players' actions.¹ Indeed Castagna and Abt concluded from their results that elite football referees varied their general coverage probably according to the technical and tactical strategies and/or fitness levels of the teams.⁷ However, Mohr et al. reported large seasonal changes in the fitness levels and physical performance of players during matches, with the players performing more HIR in matches towards the end of the season.⁸ The authors concluded that the players improved their physical performance during the season. An increased HIR by players during a match could in turn impact upon the physical demands imposed on the referee.

Krstrup and Bangsbo reported that a period of intense intermittent training improved referees' physical match performance during a game through an increased amount of HIR during the second half of matches. However, it may be that, as the matches were performed during the second part of the season, the overall match intensity was greater and this in turn impacted upon the referees work rate. Nevertheless, the influence of the intensity at which the players play the match, and also the seasonal variations in match intensity have yet to be examined in terms of their impact upon the physical match performance of the referee.

For these reasons, the aims of the present study were: 1, to examine if referees' physical match performances during the first half of a competitive match influence second half match activities; 2, to examine the relationship between the activity profile of the referee with that of the players; 3, to examine variations in match running performance across an entire competitive season; 4, to determine match-to-match variations of match analysis data.

Methods

Subjects and football match data

Objective measures of physical match performance were systematically collected from 19 fulltime professional English Football Association (FA) Premier League referees (age: 40.1 ± 4.9 years and mass: 84.1 ± 8.9 kg). The data were collected on matches during the 2004–2005 English season with a total

of 254 match observations (median of 13 matches, range 9–24 for each referee).

Match analysis data

Each match was monitored using a computerised, semi-automatic video match analysis image recognition system (data were supplied by ProZone®, Leeds, England and published with formal permission of the company). This method uses eight cameras to record the positional data of every player and official, every 1/10th of a second during a match. The cameras were positioned on the roof of the stadiums with one camera being situated in each of the stadium's four corners. The other four cameras were equally positioned along the side of the stadium. All the cameras were positioned to overlap to ensure that every movement on the pitch was captured. The data captured were then systematically analysed using proprietary software to provide an interactive coaching and analysis tool that provided comprehensive data on each individual, including how much distance each player and official covered and at what speeds they covered it. Recent findings have demonstrated that such commercial semi-automated, video match analysis image recognition systems show high accuracy and reliability, suggesting their validity for the quantification of match-related physical activities in football for practical and research purposes.⁹

The objective measures of match running performance selected for the analysis were: 1, TD; 2, HIR distance (running speed > 5.5 m/s (>19.8 km h⁻¹); 3, mean distance from infringements (i.e., fouls (DI)).

Criteria of data classification for comparisons

Effect of activity in the first half on the second half

The TD and HIR data set ($n = 254$) were ranked and divided into three subsets ('High', 'Medium' and 'Low') according to score ranks on the basis of the TD and HIR covered in the first half. Comparisons were then made to examine if the TD covered during the first half influenced the TD covered in the second half. The same procedure and analysis were also applied using HIR distance.

Activity profile in relation to the players

To assess the relationship between the referees' physical match performances with those of the players in 212 match observations, the mean HIR undertaken by the players who completed the entire match from both teams was calculated. This

was done to avoid any indirect comparisons being made with regard to playing time, and ultimately HIR between the referees who all completed the entire match with those players who did not complete the match. In terms of between-half differences in activity profile, the percentage differences in HIR were also calculated for both the referees and the players in attempt to examine any possible relationship between the changes in activity profile across halves.

Seasonal variation in match activity profiles

The seasonal variations were investigated comparing the mean TD, HIR, and DI of 15 referees during three periods of the competitive season: Start (57 matches from September to November), Mid (73 matches from December to February) and End (59 matches from March to May). For each period each referee was studied a minimum of two and a maximum of 10 times, with the mean value for each of the selected parameters of match running performance used for the analysis.

Match-to-match variations

Match-to-match variations of TD and HIR were determined from 88 matches played within a 7-day period across the competitive season from a total of 16 referees (1–8 comparisons each, with an average of 3 comparisons per referee).

Statistical analyses

Data are presented as the mean \pm standard deviation (S.D.). Before using parametric statistical test procedures, the assumptions of normality and sphericity were verified. Statistical significance was set at $p < 0.05$.

Effect of activity of the first half on the second half

Paired *t*-tests were used to compare between-half differences in TD and HIR. The relationships between first half TD or HIR and change scores calculated between the first and second half TD or HIR were examined using Pearson's product moment correlations. A factorial ANOVA with repeated measures (3×2 design) was used on each dependent variable to examine further the relationship between the amount of physical activity completed in the first half with the subsequent second half physical performance measures. The independent variables included one between-subject factor (amount of first half match distance) with three levels (High, Mid and Low), and one within-subjects factor (time) with two levels (first and second half). ANOVAs were used to verify the

null hypotheses of no differences between halves across the three groups (amount of first half match distance \times time interaction). When a significant F value was found Bonferroni's post hoc tests were applied. Effect sizes (η^2) were also calculated and values of 0.01, 0.06 and above 0.15 were considered small, medium and large, respectively.¹⁰

Relationship between referees and players activity profiles

The relationship between the mean match HIR, and also between-half percentage differences in HIR, of the referees and the players were examined using Pearson's product moment correlation. The relationship between the players' HIR and the referees' mean DI was also examined using Pearson's product moment correlation.

Effect of seasonal variation on match activity data

A repeated measures ANOVA was used to examine the seasonal variations (independent variable) in match running performance (dependent variables). Effect sizes (η^2) were again calculated.

Match-to-match variations

Paired t -tests were used to compare differences in Match 1 [M1] and Match 2 [M2] for TD and HIR. Match-to-match variations of match distances were determined using typical error expressed as coefficient of variation (CV).¹¹

Results

Effects of first half activity

The referees covered significantly less TD during the second half of the match (5790 ± 416 m versus 5832 ± 389 m, $p = 0.04$), whereas HIR was consistent between halves (391 ± 139 m versus 396 ± 142 m, $p = 0.54$). Correlation analysis revealed a significant first half versus second half relationship for TD ($r = 0.47$, $p < 0.001$, $n = 254$) and HIR ($r = 0.52$, $p < 0.0001$, $n = 254$). Also, significant negative correlations between first half physical activity and absolute changes between the first and second half for TD ($r = -0.32$, $p < 0.001$, $n = 254$) and HIR ($r = -0.48$, $p < 0.0001$, $n = 254$). When the TD and HIR data set was divided into three subsets ('High', 'Medium' and 'Low') according to score rankings on the basis of the TD and HIR covered in the first half significant interactions were found for the first half distance \times time (Fig. 1). Specifically, referees who covered less HIR in the first half

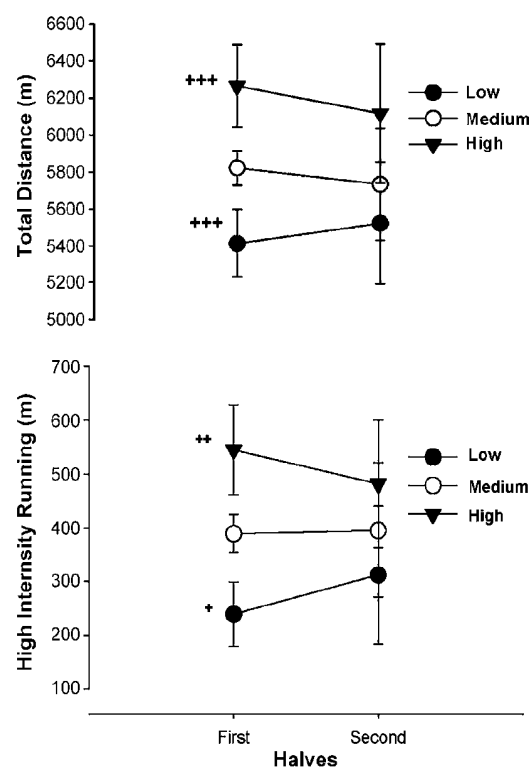


Figure 1 Effects of first half activity on second half space coverage in Premier League soccer referees. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

(238 ± 59 m, range: 74–323 m) covered more HIR in the second half (+30%), while referees who covered more HIR distance in the first half (542 ± 83 m, range: 448–835 m) decreased the amount of HIR in the second half by 11%. Similarly, referees who covered less TD in the first half (5411 ± 184 m, range: 4803–5660 m) covered more TD in the second half (2%), while referees who covered higher TD distance in the first half (6258 ± 224 m, range: 5980–6915 m) decreased the TD in the second half by 2%. Effect size for TD was medium ($\eta^2 = 0.12$) and the effect size for HIR was large ($\eta^2 = 0.16$).

Relations between referees and players

As can be seen from Fig. 2, the referees' HIR correlated moderately with the players' HIR ($r = 0.43$, $p < 0.0001$, $n = 212$). The percentage change in the referees' HIR between the first half and the second half also correlated moderately with the percentage change in the players' between-half HIR ($r = 0.44$, $p < 0.0001$, $n = 212$). There was no significant relationship between the players' HIR and the referees' DI ($r = 0.02$, $p > 0.05$, $n = 212$). Also, the referees' DI showed no significant correlation with TD ($r = -0.22$, $p > 0.05$, $n = 212$) and HIR ($r = -0.22$, $p > 0.05$, $n = 212$).

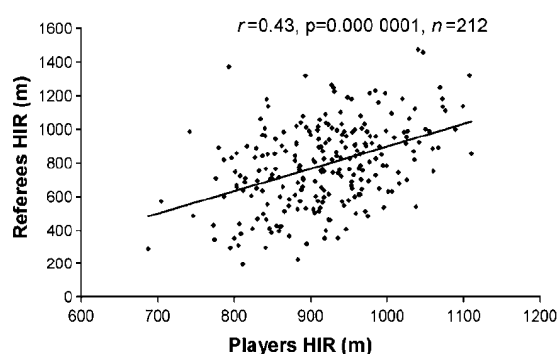


Figure 2 The relationship between the amounts of HIR covered by the PL referees and PL players $n=212$, $r=0.43$, $p<0.001$.

Seasonal variations

Seasonal variations data are shown in Table 1. No significant Start–Mid–End differences were detected for the any of the match activities variables considered in this study.

Match-to-match variations

Typical error, expressed as a CV, was 15 and 47% for TD and HIR, respectively. The CVs of the first- and second half TD were 3 and 16%, respectively. HIR showed a CV of 39% for first half values and a slightly lower value for the second half data, 35%.

No match \times half effect was observed for total distance ($F_{1, 82}=0.41$, $p=0.59$) and HIR ($F_{1, 82}=0.002$, $p=0.96$). No significant differences were observed between TD covered in Matches 1 and 2 ($11,666 \pm 660$ and $11,611 \pm 775$ m, respectively, $p=0.49$). However, referees covered significantly less distance at HIR during the second match (837 ± 239 and 761 ± 220 m,

$p=0.03$, respectively). Effect size for HIR was large ($\eta^2=0.33$).

Discussion

This study demonstrated for the first time that football referees' physical match performance is related to the overall match intensity as indicated by the amount of HIR performed by the players. Also, the second half TD and HIR performed by the referees are related to the amount of physical activities completed in the first half. No seasonal variations in the referees' physical match performances were found.

The present study reported a mean match TD covered during Premier League matches to be $11,622 \pm 739$ m. This TD is higher than the distances reported within the recent scientific literatures.^{2,3} The differences in TD between the present and previous studies might be because of differences in the fitness of top-level referees over recent years, as the referees involved in the present study were full-time professionals, or possibly differences in the technologies used to measure the distances covered. However, the type of computerised, semi-automatic video match analysis image recognition system used in the present study has a reliability of 3.2% for HIR and 1.0% for TD,⁹ which compares favourably with recently reported reliability checks on the global positioning system (GPS) and also the more traditional computer-based tracking systems.¹² However, more studies are required to validate these commercial match analysis systems fully.

In the present study, 787 ± 245 m were covered during HIR. In comparison, D'Ottavio and Castagna²

Table 1 Seasonal variations in match activities for the PL soccer referees ($n=189$)

Match variable	Start	Mid	End
Match distance (m)	11557	11703	11703
S.D.	696	763	683
First vs. second half match distance	−34.0	66.5	43.6
S.D.	157.6	215.8	195.2
HIR (m)	798.2	784.5	815.7
S.D.	228.1	235.0	200.1
First vs. second half HIR (m)	−40.8	5.41	4.01
S.D.	99.6	57.4	100.9
Players HIR (m)	903.4	927.0	915.0
S.D.	83.8	74.8	75.0
Distance from infringements (m)	14.4	14.1	14.0
S.D.	1.1	1.2	0.8

reported a mean HIR distance of 1973 ± 623 m, while Krstrup and Bangsbo³ reported 580 ± 20 m of match HIR for Danish referees. Substantial differences in the methodologies the authors used to classify HIR make direct comparisons between the results of the present study and those previously reported within the literature not possible. Given the high variability with regard to the inter-referee HIR in this study, and also within the previous literatures,^{1,2,7} in future it may be better to express HIR distance relative to individual referees' maximal speed and use similar running speed ranges. This would permit future studies to make a more consistent judgement on the inter-referee difference in match HIR and ultimately permit a more accurate analysis of the referees' physical match performances. Match heart rate responses may also be used as a measure of inter-referee physical match performances as they have been reported to be a valid measure of global match intensity in soccer referees.¹³ Weston et al.¹³ reported mean match heart rates of $83.6 \pm 2.6\%$ HR_{max} for English Premier League referees, a figure consistent with the values reported for Danish soccer referees³ and elite European referees.¹⁴

One of the main findings of the present study was the impact of the referees' physical match performance in the first half upon their physical performance in the second half. In terms of the mean half total distance the referees covered significantly less distance during the second half of the match (5790 ± 416 m versus 5832 ± 389 m, $p = 0.04$). This finding is consistent with Catterall et al.¹ and D'Ottavio and Castagna² who both reported less distance covered by the referees in the second half when compared to the first half. Given that the difference was only 48 m, which equates to less than 1% of the TD covered, the physiological significance of the difference is questionable. However, the decrease reported in the referees who completed a high amount of physical activity in the first half is higher (148 m) and, probably, of more significance in relation to fatigue. No significant difference was observed between first and second half HIR (391 ± 139 m versus 396 ± 142 m, $p = 0.54$). This is line with the findings reported by D'Ottavio and Castagna,² whereas Krstrup and Bangsbo³ reported that the amount of HIR performed by referees decreased during the second half.

In an attempt to further explore the relationships between the referees' first half activity upon second half physical performance, the TD and HIR data were divided into three subsets ('High', 'Medium' and 'Low') according to score ranks on the basis of the TD and HIR covered in the first half. Ref-

erees whose TD was classified as high in the first half significantly decreased their TD in the second half, whereas referees whose TD was classified as low during the first half significantly increased their TD in the second half. The same pattern of results was also observed for HIR. For both TD and HIR, those referees who covered a medium distance during the first half displayed no significant change in TD and HIR in the second half. Castagna and Abt⁷ reported that experienced referees may moderate their competitive behaviour during a match and that this 'sparing behaviour' may be adopted by the referees to avoid fatigue. The results of the present study may support this observation as those referees who covered a smaller TD and HIR in the first half had the available physical capacity to be able to increase their physical performances during the second half.

Despite statistical significance, the 2% change in TD found during the second half in those referees who covered less or more TD during the first half may be considered of relatively minor physiological importance. However, the 30% increment and 11% decrement in HIR observed when referees covered less or more HIR in the first half, respectively, may be considered as relevant to fatigue determinism. In this regard, Krstrup and Bangsbo³ stated that a reduced distance covered by HIR suggested that referees were fatigued towards the end of the game.

Furthermore, the effect size was large for HIR and medium for TD. This analysis was also confirmed by the significant higher negative correlations between first half HIR and absolute change in HIR between the first and second half compared to the association found for TD ($r = -0.48$ versus $r = -0.32$; $p < 0.05$). This seems to suggest that HIR is a better indicator of match physical demands than TD, especially in relation to the development of fatigue.⁷ Castagna and Abt have already reported that HIR may be considered as a more reliable variable for monitoring the tactical strategy of the elite-level referee than TD.⁷

Mohr et al.⁸ highlighted that there are natural variations in the intensity of games due to tactical or psychological factors and Burgess et al.¹⁵ suggested from their findings on the movement demands of national players in Australia that players may indeed selectively spare energy for more crucial high-intensity efforts. With regard to the referees, Catterall et al. stated that the overall pattern of a referee's match activity was acyclical but varied in parallel with the players' actions.¹ Indeed, Castagna and Abt concluded from their results that elite soccer referees varied their general coverage probably according to the techni-

cal and tactical strategies and/or fitness levels of the teams.⁷ However, the supposed relationship between the activity profiles of the referee with the overall match intensity of the players was not based on quantitative data with regard to its influence on the referees' physical match performance. Using the amount of HIR performed by the players as a measure of overall match intensity, as this is most reflective of the physical demands of a match,³ a weak but significant correlation was found in the present study between the activity of profile of the players and that of the referees. Also, the percent change in the referee HIR between the first half and the second half was also significantly correlated with the percent change in the players between-half HIR. However, the significant correlations may have been also influenced by the large sample size ($n=212$) and data variability.

A limitation of this study was that, given the strength of our results, the physical match performances of the referees, determined during actual match analysis, are probably influenced by other factors not measured in the present study. However, our data do demonstrate for the first time that the referees' physical match performances are related, in part, to the physical performances of the players and thus the overall intensity of the match itself. Therefore, the results would appear to have implications when assessing a referee's match performance in that there may be a need to assess the overall match intensity prior to examining the physical match performance of the referee to gain a further insight into the pattern of the match. Also, referees' reduced physical performances during the second half of a match require further examination as to whether they are directly attributable to fatigue or merely a consequence of a slower tempo of the match in the second half.

The most important aspect of refereeing is the decision-making process, because throughout the game a referee has to make correct decisions under time constraints.¹⁴ Referees are required to keep up with play at all times and ensure optimal viewing positions.¹⁶ However, in the present study there was no relationship between the referees' physical match performances and DI, a variable commonly considered as relevant to the decision-making process in football refereeing.¹⁷ Therefore, it is recommended that further studies examining the relationship between the referees' physical and technical performances are undertaken because previous research has reported that the intensity of match play in soccer referees corresponds to an exercise intensity that has been shown to cause decrements in cognitive and psychomotor behaviour.³

Previous research has reported that important measures of football physical match performance, such as TD and HIR, increased during a competitive season.⁸ Mohr et al. observed large seasonal changes in physical match performances with soccer players demonstrating an improved physical match performance, through a greater HIR distance, towards the end of the season.⁸ However, in the present study, no seasonal variations were found for any of the referees' match activity variables and also the players' HIR. As we found a significant relationship between the physical performance of referees and soccer players and the players' physical match performances in the present study did not change as the season progressed, this may in part explain why no seasonal variations were observed in the referees' TD and HIR. Also, Mohr et al. suggested that the change in HIR observed in soccer players was probably because of the lack of time for physical training during the week, since at the start and middle of the competitive season there are more matches a week for Champions League and National Cup competitions.⁸ This could result in a decrease of players' physical fitness and therefore the ability to perform HIR during the matches. This probably did not happen in the present study as not all the referees analysed were involved in National Cup or Champions League matches.

The match-to-match variability for TD and HIR were much higher than those reported in previous research.^{8,14} with values of 15 and 47%, for TD and HIR, respectively. Also match-to-match variations calculated on 16 referees within a 7-day period during the competitive season reported no between-match variations in TD. However, despite there being no match \times half effect for TD and HIR, referees' HIR was significantly less during the second match. These results demonstrate that a referee's match HIR is more variable than TD. The players' HIR was not examined alongside the referees' match-to-match variability data as different teams and players were involved in the referees' matches and therefore no measure of match-to-match variation was possible for the players.

However, from the results presented in the current study, it is difficult to conclude whether the reduced HIR was due to fatigue, referees' tactical strategies or merely a slower tempo of match on behalf of the players, as has been previously reported.^{4,8} These large match-to-match variations, particularly for HIR, would appear to have implications for the appropriate planning of referee fixtures and also the need for fitness development to enable the referees to be sufficiently able to cope with the physical demands of their matches.

Conclusion

This study reported for the first time that football referees' physical match performance is related to overall match intensity, as indicated by the amount of HIR performed by the players. Therefore it is suggested that, when assessing a referee's physical match performance, there is a need to assess the overall match intensity prior to examining the activity profile of the referee in order to gain a further insight into the pattern of the match. Also, our results suggest that, the higher the amount of HIR completed during the first half, the lower the ability to sustain high-intensity physical activity during the second half, possibly as a consequence of accumulated fatigue. No seasonal variations in the referees' physical match performances were found. Therefore, further studies are required to measure both physiological and match performance changes during the season in football referees.

Practical implications

- Football referees' physical match performances are partly related to those of the players.
- Assessment of overall match intensity should be undertaken prior to examining the physical match performance of the referee.
- Match-to-match variations within a 7-day period demonstrate that referees' high-intensity running is significantly reduced during the second match.
- Appropriate planning of referee fixtures is required to enable the referees to recover sufficiently and in turn be able to cope with the physical demands of all of their matches.

References

1. Catterall C, Reilly T, Atkinson G, et al. Analysis of work rate and heart rates of association football referees. *Br J Sports Med* 1993;27:153–6.
2. D'Ottavio S, Castagna C. Analysis of match activities in elite soccer referees during actual match play. *J Strength Cond Res* 2001;15(2):167–71.
3. Krustup P, Bangsbo J. Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. *J Sports Sci* 2001;19:881–91.
4. Bangsbo J, Norregaard L, Thorso F. Activity profile of competition soccer. *Can J Sport Sci* 1991;16(2):110–6.
5. Stølen T, Chamari K, Castagna C, et al. Physiology of soccer: an update. *Sports Med* 2005;35(6):501–36.
6. Johnston L, McNaughton L. The physiological requirements of soccer refereeing. *Aust J Sci Med Sport* 1994;26(3/4):67–72.
7. Castagna C, Abt G. Inter-match variation of match activity in elite Italian soccer referees. *J Strength Cond Res* 2003;17(2):388–92.
8. Mohr M, Krustup P, Bangsbo J. Match performance of high-standard soccer players with special reference to development of fatigue. *J Sports Sci* 2003;21:519–28.
9. Impellizzeri FM, Sassi A, Rampinini E. Accuracy and reliability of a commercial video-computerized, semi-automatic, soccer-match analysis system: preliminary results. In: Hoppeler H, Reilly T, Tsolakidis E, Gfeller L, Cologne KS, editors. *11th Annual Conference of the European College of Sport Science*. Lausanne, Switzerland: Sportverlag Strauss; 2006. p. 319–20.
10. Cohen J. *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
11. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med* 2000;30:1–15.
12. Edgecomb SJ, Norton KI. Comparison of global positioning and computer-based tracking systems for measuring player movement distance during Australian Football. *J Sci Med Sport* 2006;9(1–2):25–32.
13. Weston M, Bird S, Helsen W, et al. The effect of match standard and referee experience on the objective and subjective match workload of English Premier League referees. *J Sci Med Sport* 2006;9(3):256–62.
14. Helsen W, Bultynck JB. Physical and perceptual-cognitive demands of top-class refereeing in association football. *J Sports Sci* 2004;22(2):179–89.
15. Burgess DJ, Naughton G, Norton KI. Profile of movement demands of national football players in Australia. *J Sci Med Sport* 2006;9(4):256–62.
16. Weston M, Helsen W, MacMahon C, et al. The impact of specific high-intensity training sessions on football referees' fitness levels. *Am J Sports Med* 2004;32(1 Suppl.):54S–61S.
17. Verheijen R, Oudejans R, Beek PJ, et al. Factors affecting decision making of soccer referees. In: Spinks W, Reilly T, Murphy A, editors. *Fourth world congress of science and football*. Sydney, Australia: University of Technology; 1999. p. 28.

Available online at www.sciencedirect.com



ScienceDirect

4.1

STUDY IV

Castagna, C., Abt, G., D'Ottavio, S. & Weston, M. (2005) Age related effects on fitness performance in elite-level soccer referees. *Journal of Strength and Conditioning Research*, 19 (4), 785-790.

AGE-RELATED EFFECTS ON FITNESS PERFORMANCE IN ELITE-LEVEL SOCCER REFEREES


CARLO CASTAGNA,¹ GRANT ABT,² STEFANO D'OTTAVIO,¹ AND MATTHEW WESTON³

¹School of Sport and Exercise Sciences, Faculty of Medicine and Surgery, University of Rome Tor Vergata, Rome, Italy; ²St. Martins' College, Lancaster, UK; ³The Football Association Premier League, Connaught Place, London, UK.

ABSTRACT. Castagna, C., G. Abt, S. D'Ottavio, and M. Weston. Age-related effects on fitness performance in elite-level soccer referees. *J. Strength Cond. Res.* 19(4):785–790. 2005.—Elite soccer referees usually achieve the peak of their officiating careers at an average age that is considerably older than that observed in competitive matched soccer players. As ageing has been reported to negatively affect physical performance, the aim of this study was to investigate the effect of ageing on fitness performance in elite-level soccer referees. Thirty-six elite-level soccer referees were grouped into young (Y, $n = 12$), average (A, $n = 14$), and old (O, $n = 10$) groups, according to their age and observed for field test performance (countermovement jump [CMJ], 50-m and 200-m sprints, and 12-minute running for distance). Results showed a significant age effect on CMJ ($r = -0.52$, $p < 0.001$), 200 m ($r = 0.51$, $p < 0.001$), and 12-minute time trial running ($r = -0.52$, $p < 0.001$). Y jumped higher than A and O groups ($p < 0.05$) and were faster than O over 200 m ($p < 0.05$). No group effect was observed for 12-minute run and 50-m performance ($p > 0.05$), respectively. Sixteen of the 36 referees were further examined for selected physiological variables and grouped into 2 equal ($n = 8$) age groups (young and old, Y1 and O1, respectively). $\dot{V}O_{2\max}$ was higher in Y1 ($p < 0.05$), but O1 attained performances similar to Y1 running at selected blood lactate speeds (4 mmol L^{-1} , $p > 0.05$). Although older referees demonstrated acceptable fitness levels, younger officials should ensure that they develop appropriate levels of aerobic and anaerobic fitness to be able to match the demands placed upon them while refereeing throughout their careers. To promote this, fitness test standards should be age related.

KEY WORDS. association football, aerobic fitness, fitness tests, ageing process, lactate threshold

INTRODUCTION

occer refereeing has been reported to elicit a cardiovascular strain that on average attains 85–95% of the estimated and individual maximal heart rate during competitive matches (13, 15, 22, 23, 28). Direct metabolic assessment during friendly matches has shown that referees officiated on average at 68% of their $\dot{V}O_{2\max}$ (17). Additionally the studies that have addressed the match aspects of soccer refereeing have demonstrated that during a competitive match a referee can cover a mean distance of 11.5 km, with ranges from 9 to 14 km (2, 13, 15, 16, 22, 23). Although from the physiological point of view the physical stress imposed on the elite soccer referee could resemble that found in soccer players playing in the mid-field (4, 16), several aspects of a referee's performance distinguishes it from that of a player's performance. For instance, officials are not involved with the ball, and cannot be substituted during the match. Furthermore, compared to the soccer players that they normally officiate,

referees have only recently (and in limited numbers) become full-time professionals.

Another relevant aspect of soccer refereeing is the existing age difference between soccer players and soccer referees. For example, Bangsbo (3) reported that the average age of players competing in the highest Danish league during the 1991–92 season was 24 years. In contrast, the average age of referees currently officiating in the Italian soccer first division (serie A and B, 2004 season) is 35 years.

The difference in the average age of players and referees may exist because experience is considered among the international refereeing governing bodies (Federation Internationale de Football Association, FIFA, and Union Europenne de Football Association, UEFA) as a fundamental prerequisite to officiate at the elite level (18). Paradoxically, elite soccer referees reach their best performance level at an average age when most soccer players have retired from competition. Usually elite-level soccer referees reach their “gold-age” career level above 40 years of age. Demonstration of that comes from the recent 2002 FIFA Federation World Cup Finals in which the average age of the super-elite-level soccer referees that officiated competitions from the quarterfinals on was 41 ± 4 years ($n = 8$, [www.fifaworldcup.yahoo.com; accessed July 25, 2002]).

Ageing is a multifactorial process that alters structures and reduces functions of cells and tissues of many organ systems (8). Consequently, the aging process has potentially a profound effect upon the physiological aspects of human performance. However, to our knowledge there are no scientific studies currently available examining the age-related effects on fitness performance in elite soccer referees.

More detailed information regarding this issue would be of extreme interest to those involved in the performance development and selection of elite-level officials (FIFA, UEFA, and affiliated country referees' associations) as fitness levels and particularly aerobic capacity have been demonstrated to be related to match physical and tactical performance (10, 11, 18, 20, 23). The documented pace increments observed in modern soccer (29) and, consequently, the need for an appropriate fitness level of the match official to cope with the increasing demands of the game promotes even further the urgency for gaining information on this issue.

The aim of this study was to investigate the possible age-related variations in fitness performance in a group of highly competitive soccer referees to obtain information to be used to guide selection and training prescription.

METHODS

Experimental Approach to the Problem

The research was carried out examining the performance and physiological responses of 36 elite referees at the time of investigation enrolled in the Italian Elite Soccer Referee Commission (Commissione Arbitri Nazionali, CAN). That represents the entire group of referees that each competitive year is considered to have the ability to officiate in the professional Italian soccer championships (serie A and serie B, first and second soccer division, respectively) by the Italian Soccer Referees Association (AIA).

In Italy no more than 36 referees are selected to officiate in the serie A and B championships each year. Every season those who exceed 45 years of age or are considered not to have the ability to cope with the demands of the game are dismissed. Because of the rigorous selection process and the high level of play taking place in the serie A and B championships, the referees belonging to this group may be considered representative of elite-level officials.

To test the likelihood of age-related variations in fitness performance, the referees were grouped into 3 arbitrary chosen age groups: young (Y, 31–35 years, $n = 10$), average (A, 36–39 years, $n = 14$), and old (O, 40–45 years, $n = 12$). These groups were tested using the field fitness test battery currently in use by UEFA (18). Additionally, with the aim to deepen the analysis on possible age-related effects on fitness performance, 16 referees were randomly chosen among the Y and O groups to constitute 2 subsets (8 referees in each subgroup) of referees, representing the younger and older officials and then submitted to physiological assessment. No significant ($p > 0.05$) differences were observed in the time spent at the selected training heart-rate zones considered for the aerobic fitness training sessions during the pretesting period among the 3 subgroups. Volume (number of repetitions and sprint distance) of anaerobic work performed by the 3 subgroups of referees showed no significant ($p > 0.05$) differences across the age groups. Whole-group ($n = 36$) regression analyses were performed to determine possible age-related variations in fitness performance.

Subjects

Participants consisted of 36 male soccer referees (age 37 ± 3 years, height 181 ± 5 cm, body mass 77.5 ± 5.1 kg) belonging to AIA and officiating in the serie A. The officials who volunteered for the present study possessed at least 10 years of refereeing experience and some of them were ranked in the UEFA list. All the referees were part-time professionals and trained at least 3 times a week, following a supervised training program aimed at improving aerobic, anaerobic, and flexibility fitness levels (Table 1).

Training loads addressing the various aspects of fitness were equated among the referees that volunteered in this study during the 2 months that preceded all the fitness assessments. Supervisors had to strictly stick to the training program developed by the third and first authors of this study. Every week the training loads provided to each referee were checked by those researchers for uniformity. Training intensity during the aerobic workloads were assessed using heart rate monitors (Polar, Finland). The referees participating in this study were in-

TABLE 1. Weekly training schedule used by referees during the 2 months preceding testing sessions (preseason training period). Referees trained during the pretest periods a minimum of 3 times a week. Speed training loads were adjusted during the training period to promote the adaptive responses.

Monday	Tuesday	Friday
10-min warm-up 10 agility training 30–60 min continuous running at 70–80% HR _{max} *	10-min warm-up 15 agility training Sprint training over short distances 10 × 30–80 m sprints with full recovery	10 min warm-up 10 agility training 3–4 × 1000 m at 90–95% HR _{max} with 3-min active recovery
10-min cool-down; 5 running + 5-min static stretching	10-min cool-down; 5 running + 5-min static stretching	10-min cool-down; 5 running + 5-min static stretching

* HR_{max} = maximal heart rate.

involved in refereeing an average of 2 times per month, from September to June (serie A and B competitive season).

Written consent from all subjects was obtained before the commencement of the study, after a brief but detailed explanation as to the nature of the research. The research design received formal approval from the ethics committee of Tor Vergata University, and AIA. Each test session was performed with well-rested subjects and usually all test bouts were completed during the first hours of the afternoon, to match the period of the day when most serie A matches are played.

Fitness Tests

All the referees ($n = 36$) were tested using the fitness test battery adopted by UEFA to assess the eligibility to officiate at the international level. The referees were tested for the 50-m and 200-m sprints and 12-minute run. Additionally, lower limb explosive strength performance was assessed using a countermovement vertical jump (CMJ) according to the methods and procedures suggested by Bosco et al. (1983) using a switch-mat connected to a computer (Ergojump, Globus, Codognè, Treviso, Italy). The 50-m sprint performance was assessed using photocell beams (Globus, Godognè). The 200-m and 12-minute run tests for distance performance were timed using stopwatches (Casio, Japan). Mean air temperature and relative humidity during testing sessions were 22°C and $44 \pm 9\%$, respectively. The testing order was CMJ, 50-m sprint, 200-m sprint, and 12-minute run test. After each test bout the referees were prescribed a recovery of 2 minutes of semiactive rest between the CMJ and 50-m sprint and between 50-m and the 200-m sprints. At least 8 minutes of self-administered rest were allowed between the 200-m sprint and the start of the 12-minute time trial running test. This recovery period was chosen as it proved to be the minimum requested for not affecting endurance performance in elite-level soccer referees (Castagna and D'Ottavio 1994, unpublished data). All test assessments were implemented by qualified fitness trainers on behalf of the AIA serie A and B CAN and performed on a tartan athletic track. Field testing was performed during the annual preseason training camp of the 1997–98 competitive season. Reliability investigations performed with Italian elite-level soccer referees showed in-

traclass correlation coefficients ranging from of 0.92 to 0.94 for the field tests used in this study ($p < 0.01$, unpublished data). The reliability (expressed as percentage coefficient of variation, CV%) of these methods were 2.1% for CMJ, 0.9% for 50-m sprint, 1.1% for 200-m sprint, and 1.7% for 12-minute run test.

Physiological Testing

Participants in the substudy were 16 soccer referees that were selected from the 36 who volunteered for the main study. They were selected within the young and old groups, 8 participants per group. These 2 paired subgroups were divided according to their age to represent the younger (Y1) and older (O1) referee population. The referees that volunteered in this substudy were tested under field conditions for the assessment of their aerobic fitness. Each of the 16 referees was assessed for maximal effort ventilation ($\dot{V}_{E\max}$), peak absolute and relative to body mass $\dot{V}O_2$, peak posteffort blood lactate concentration, and maximal heart rate (HR_{\max}). Submaximal aerobic performance was also assessed by calculating the running speed attained at a blood lactate concentration of 2 and 4 mmol L^{-1} . The velocities attained at 2 fixed blood lactate concentrations (2 and 4 mmol L^{-1} , \dot{V}_2 , and \dot{V}_4 , respectively) were selected, as these are commonly used to assess endurance performance and have also been shown to correlate with endurance performance (19). \dot{V}_2 and \dot{V}_4 were computed with interpolation when not directly found. Running economy (RE) was calculated as the $\dot{V}O_2$ elicited by running at 10 km h^{-1} . Expired gases were analyzed using a portable lightweight gas analyzer (K2, Cosmed, Rome, Italy) that enabled $\dot{V}O_2$, VE, and HR data collection. Gas and HR sampling were performed every 10 and 5 seconds, respectively. Variables of interests were obtained with subjects performing a progressive multistage protocol conducted under field conditions with the K2 technology. Validity and reliability of the K2 technology have been previously described by Lucia et al. (24). Reliability and validity of the test protocol used to assess the physiological variables of interest are reported elsewhere (10, 11).

After a brief warm-up carried out on an individual basis, subjects ran at 8 km h^{-1} for 2 minutes. Thereafter, the speed was increased every 2 minutes by 2 km h^{-1} until volitional exhaustion. At the end of each stage subjects were stopped to allow blood sampling (30 seconds). Earlobe blood samples were drawn immediately before the start of the test protocol and at the end of each 2-minute stage. Blood samples were collected in a capillary tube containing a fluoride/heparin/nitrite mixture. Blood lactate concentrations were determined by an amperometric enzymatic method (Roche Lactate Analyser 640, Paris, France). The subjects performed the test running on an athletic track marked with cones every 20 m. The pace was set with the help of an audio player broadcasting prerecorded beeps of an audiocassette. The beep indicated the moment when the subjects had to pass near a cone to maintain a constant speed. The test protocol was a modified version of the Montreal Track Test (12), with the same stage duration but faster speed increments to reduce test time. The minimum speed that elicited maximal $\dot{V}O_2$ values during the last stages of the multistage track test were considered as representative of the speed at $\dot{V}O_{2\max}$ (6). Subjects who are tested with this protocol usually become exhausted within 10–12 minutes of ex-

TABLE 2. Anthropometrical characteristics of young, average, and old groups.

Variables	Group		
	Young ($n = 12$)	Average ($n = 14$)	Old ($n = 10$)
Age (y)	33.5 \pm 2.5*‡	37.5 \pm 1.0†‡	42 \pm 1*†
Height (cm)	181 \pm 2.5	182 \pm 1.0	180 \pm 2.5
Body mass (kg)	77.5 \pm 2.1	76.8 \pm 1.8	77.0 \pm 2.1

* = significantly different from average group, $p < 0.05$. † = significantly different from young group, $p < 0.05$. ‡ = significantly different from old group, $p < 0.05$.

ercise (12). Field tests were used because referees felt more motivated under field conditions than under laboratory conditions, thus producing more reliable results.

Statistical Analyses

Data are presented as mean and SD. Differences between 2 groups were tested for significance using *t*-test for unpaired samples. Comparisons between more than 2 means were performed using one-way analysis of variance. Post hoc testing was carried out using Tukey's test. Relations between variables of interest were detected using Pearson's correlation coefficient (r). Estimations were performed calculating bivariate regression equations. Significance was set for the calculations at 5% ($p \leq 0.05$).

RESULTS

Fitness Tests

Entire group ($n = 36$). Average 50-m sprint time was 7.00 \pm 0.32 seconds. During the 12-minute run test for distance the referees covered 2,838 \pm 179 m. The 200-m sprint and CMJ performances were 29.89 \pm 1.66 seconds and 34 \pm 3 cm, respectively.

Subgroups (Y, A, O) and physiological study. Age, height, and body mass of the of Y, A, and O groups are presented in Table 2. In the CMJ test the Y, A, and O groups scored 36.31 \pm 3.25, 33.00 \pm 2.93, and 32.6 \pm 3.02 cm, respectively. Significant performance differences were found between O and Y ($p < 0.03$), and between Y and A groups ($p < 0.03$). Running 12 minutes for distance the Y, A, and O groups scored 2,912 \pm 175, 2,846 \pm 189, and 2,740 \pm 130 m, respectively. Group differences in the endurance test performance were not significant ($p > 0.05$). Sprinting over 50 m the Y group scored 6.89 \pm 0.30 seconds, whereas subjects in the A and O groups covered the distance in 7.07 \pm 0.33 and 7.07 \pm 0.31 seconds, respectively. Short-sprint performances were not significantly different among the 3 groups ($p > 0.05$). Differences in the 200-m sprint test scores were found between the O and Y groups ($p < 0.05$, 30.73 \pm 1.36 and 28.95 \pm 1.36 seconds, respectively). In the 200-m sprint test the referees belonging to the A group scored 29.89 \pm 1.80 seconds.

To detect possible age-related effects on field test performances, regression analyses were performed plotting test scores against age and considering all the referees as single group ($n = 36$). The 12-minute run test and CMJ performances showed negative correlations with age ($r = -0.52$ for both conditions, $p < 0.01$). A positive correlation of the same absolute magnitude level was found between 200-m sprint performance and age ($r = 0.51$, $p < 0.001$). No significant age-related effects were detected for the

TABLE 3. Physiological and anthropometrical variables assessed in the older and younger groups.

Variables	Group	
	Older <i>n</i> = 8	Younger <i>n</i> = 8
Age (y)	42.1 ± 1.2	33.1 ± 1.8*
Height (cm)	184 ± 5	182 ± 2
Body mass (kg)	79.5 ± 7.6	77.0 ± 7.5
RE† (ml kg ⁻¹ min ⁻¹)	37.6 ± 2.8	40.5 ± 5.7*
VE _{max} (l min ⁻¹)	142 ± 23	158 ± 3*
VO ₂ (l min ⁻¹)	3.33 ± 0.56	3.95 ± 0.66*
VO ₂ (ml kg ⁻¹ min ⁻¹)	42.50 ± 4.46	52.10 ± 7.36*
V ₂ (km h ⁻¹)	10.02 ± 2.14	9.50 ± 1.66*
V ₄ (km h ⁻¹)	12.62 ± 1.61	12.75 ± 1.08
Speed at VO _{2max} [km h ⁻¹)	15.5 ± 1.0	16.5 ± 1.9
HR _{max} (beat min ⁻¹)	176 ± 5	189 ± 4*
L _{max} (mmol l ⁻¹)	10.08 ± 2.23	11.57 ± 2.30*
12-min run test (m)	2,873 ± 149	2,934 ± 136

* = significantly different from older group, $p < 0.05$.

† RE = running economy; VE_{max} = maximal effort ventilation; V₂ = velocity attained at blood lactate concentration of 2 mmol L⁻¹; V₄ = velocity attained at blood lactate concentration of 4 mmol L⁻¹; HR_{max} = maximal heart rate; L_{max} = peak post-effort blood lactate concentration.

50-m sprint performance. Results of the selected variables considered in the physiological study are showed in Table 3.

DISCUSSION

The results of the present study show that there was no significant difference in endurance performance between age groups when assessed as distance covered during the 12-minute time trial running test. This finding contrasts with the results reported by other studies reporting profound age-related variations in endurance, with the age-related decrease in aerobic performance being attributed to predominantly central factors, i.e., a reduced responsiveness to sympathetic stimulation causing a decrease in HR_{max} with advancing age, as opposed to peripheral factors (27). Although no endurance performance difference occurred across age groups, regression analysis revealed the existence of a significant trend toward an aerobic age-related performance impairment. This result supports the idea that through regular fitness training age-related differences in physical performance may be reduced, which is consistent with Rogers et al. (26), who reported that a relatively constant level of regular vigorous endurance exercise training reduced the rate of the age-associated VO_{2max} decline from 10% per decade to ~5%.

The physiological assessments performed in this study revealed significant differences between the Y1 and O1 groups in the majority of the aerobic metabolism variables studied (see Table 3). Those referees belonging to the Y1 group had higher levels of VO_{2max} both in absolute and relative terms ($p < 0.05$), revealing strong age-related differences in maximal aerobic performance.

Interesting results were found in the age-related variation in RE. In fact, those referees in the Y1 group, despite a superior maximal oxygen uptake expressed both in absolute and relative to body mass terms, possessed a poorer running economy compared to their older counterparts. Y1 referees were significantly less economical and slower than the O1 group running at 10 km h⁻¹ and at a

speed that elicited a blood lactate of 2 mmol L⁻¹, respectively. Those between-group differences disappeared when the speed attained at 4 mmol L⁻¹ was considered.

However, when expressing RE as a percentage of their individual maximal aerobic power, Y1 and O1 ran at 77.7% and 88.4% of their VO_{2max}, respectively ($p < 0.001$). This finding reveals that O1 referees were able to match Y1 group aerobic performance through a more efficient use of their limited (compared to Y1 levels) maximal aerobic power. These interesting results may help to explain the similar responses in field test endurance performance observed by both groups.

Corroboration of this comes from the finding that 12-minute run test performance (distance covered) was positively and strongly related with V₄ ($n = 16$, $r = 0.77$, $p < 0.05$). However, relative to body mass VO_{2max} showed a similar relationship with the 12-minute run time-trial performance ($n = 16$, $r = 0.79$, $p < 0.05$). Additionally, analysis of the velocities associated with VO_{2max} revealed no significant difference between Y1 and O1 ($p > 0.05$) but significant correlation ($r = 0.83$, $p < 0.01$) was found between speed at VO_{2max} and maximal oxygen uptake.

Considering these results we may say that a well-developed submaximal aerobic performance is probably the reason why those referees belonging to the older age group perform as well as their younger counterparts in the 12-minute run test for distance. This result is particularly important, as 12-minute run test performance and speed attained at 4 mmol L⁻¹ have been proven to positively influence the physical match performance of elite-level soccer referees (11). However, the proved effect of the individual level of relative-to-body-weight maximal oxygen uptake on maximal aerobic speed and 12-minute time trial run test may urge the necessity for emphasis on VO_{2max} development. With respect to this, VO_{2max} has been shown to positively affect match performance in elite-level Italian soccer referees (9).

These results are in agreement with what has previously been found by Allen et al. (1). In their study Allen et al. (1) reported that masters endurance athletes were able to perform as well as competitive matched younger runners despite a 9% lower VO_{2max}. The ability of the older endurance athlete to work closer to their VO_{2max} was considered as a possible reason for the similarity in endurance performance (1).

The detected differences in submaximal aerobic performance between age groups in the present study lends support for the need to assess more variables than VO_{2max} alone when determining the aerobic performance of elite-level soccer referees. However, the design of the current study does not allow us to say whether the superior RE observed in the O1 group was determined by genetic or training variables.

Differences in submaximal running performance have been reported to be related to training running speeds (5). Although during the period preceding this study efforts were made to equate training volume and intensity among participants, training intensity was assessed on an individual basis using heart rate monitors. For this reason it is not possible to say whether the better RE observed in the O1 group could be due to a larger use of slower training speeds than in the Y1 group. However, as the average speed at VO_{2max} was not significantly different between the two subgroups, it seems that O1 group

possessed a well-developed across-speeds running efficiency.

CAN referees showed age-related differences only in the short (1–35 seconds), predominantly anaerobic performance domain and specifically in the explosive-strength and anaerobic endurance tests. This finding is in agreement with the results reported by Pirnay and Crielaard (25), who showed significant age-related decrements in anaerobic performance from 25 years of age onwards. When comparing the declines in aerobic and anaerobic work capacity in the elderly, Chamari et al. (14) reported that anaerobic power declines with age more rapidly than aerobic power and attributed a decrease in the type II:I muscle fiber ratio as being the primary mediator responsible for the lower anaerobic to power aerobic ratio. Differences in CMJ performance between O and Y groups are in agreement with the results reported by Bosco and Komi (7), who showed age-related loss of leg power during vertical jumping.

It could be speculated that requiring soccer refereeing 68% of $\dot{V}O_{2\max}$ (17) at 85–95% of the estimated and individual HRmax (15, 23, 28) per 90 minutes, match participation may have a training effect (21). This may be particularly true for those referees that are older, as these officials may have accumulated more competitive experience compared to younger colleagues. In this regard Bangsbo (3) reported no difference in aerobic fitness level in regular and nonregular first-team soccer players. Although match participation frequency is lower than for soccer players (1–2 and 4–6 matches per month for referees and players, respectively), we cannot exclude that in this population of moderately trained subjects, match participation played a positive role in fitness improvement and maintenance. However, this study design may not give conclusive responses in this regard. Because of the interest of this issue, further research may be quite helpful for accurate training prescription.

Although not proven to be directly related to match performance (10), modern soccer elite-level referees seem to require a good level of anaerobic fitness (17, 23). With respect to this Krstrup and Bangsbo (23) have shown that during a competitive match top-level soccer referees on average perform 150–200 high-intensity running activities interspersed with periods of low-intensity exercise. High blood lactate concentrations have been reported during the match, demonstrating that the anaerobic system may be highly taxed in a given period of the game (17, 23). These findings may suggest to fitness trainers to pay attention in implementing specific anaerobic training as age progresses in elite soccer referees. However, experienced elite-level soccer referees may overcome a lack of anaerobic fitness by being more selective in their high-intensity activities (9, 16).

PRACTICAL APPLICATIONS

Considering that impairment of physical performance can only be lessened but not probably completely avoided, national and international (FIFA and UEFA) refereeing governing bodies should make efforts to select talented referees who possess a good level of performance in the aerobic and anaerobic field tests for their age. In this regard national and international refereeing organizations might find age-related fitness standards and not age-independent criteria as currently in use (18) of some help. In fact, considering the age-related decline in 12-minute

run test performance detected in this study, it is estimated that a 30-year-old soccer referee should score 3,030 m to be able to pass the 2,700-m standard set by UEFA (18) at the age of 42, which is approximately the average age of the officials who directed the last stages of the 2002 FIFA World Cup. It is suggested that elite-level soccer referees should be able to cover in their early-stage career at least 3 km in the 12-minute running test, scoring 6.90 seconds and 28 seconds over 50-m and 200-m sprint tests, respectively (10). Referees should also possess a $\dot{V}O_{2\max}$ not lower than 50 ml $\text{kg}^{-1} \text{min}^{-1}$ when debuting at the elite level. Finally, comprehensive physiological testing is highly advisable in elite-level soccer referees to observe possible weak and strong performance points that may then be addressed by adopting appropriate training strategies. A professional status similar to the soccer players whom they officiate should also be given consideration.

REFERENCES

1. ALLEN, W.K., D.R. SEALS, B.F. HURLEY, A.A. EHSANI, AND J.M. HAGBERG. Lactate threshold and distance-running performance in young and older endurance athletes. *J. Appl. Physiol.* 58:1281–1284. 1985.
2. ASAMI, T., H. TOGARI, AND J. OHASHI. Analysis of movement patterns of referees during soccer matches. In: *Science and Football*. T. Reilly, et al., eds. London: E & FN Spon, 1988. pp. 341–345.
3. BANGSBO, J. The physiology of soccer—With special reference to intense intermittent exercise. *Acta. Physiol. Scand.* 151(Suppl. 619):1–155. 1994.
4. BANGSBO, J., L. NORREGAARD, AND F. THORSO. Activity profile of competition soccer. *Can. J. Sport Sci.* 16:110–116. 1991.
5. BERG, K. Endurance training and performance in runners: Research limitations and unanswered questions. *Sports Med.* 33: 59–73. 2003.
6. BILLAT, V., AND J.P. KORALSZTEIN. Significance of the velocity at $\dot{V}O_{2\max}$ and time to exhaustion at this velocity. *Sports Med.* 22:90–180. 1996.
7. BOSCO, C., AND P.V. KOMI. Influence of aging on the mechanical behaviour of leg extensor muscles. *Eur. J. Appl. Physiol.* 45:209–219. 1980.
8. BRUCE, R.A. Exercise, aerobic capacity and ageing—Another viewpoint. *Med. Sci. Sports Exerc.* 16:8–13. 1984.
9. CASTAGNA, C., AND G. ABT. Intermatch variation of match activity in elite Italian soccer referees. *J. Strength Cond. Res.* 17: 388–392. 2003.
10. CASTAGNA, C., G. ABT, AND S. D'OTTAVIO. Relation between fitness tests and match performance in elite Italian soccer referees. *J. Strength Cond. Res.* 16:231–235. 2002.
11. CASTAGNA, C., G. ABT, AND S. D'OTTAVIO. The relationship between selected blood lactate thresholds and match performance in elite soccer referees. *J. Strength Cond. Res.* 16:623–627. 2002.
12. CASTAGNA, C., AND S. D'OTTAVIO. Effect of maximal aerobic power on match performance in elite soccer referees. *J. Strength Cond. Res.* 15:420–425. 2001.
13. CATTERALL, C., T. REILLY, G. ATKINSON, AND A. COLDWELLS. Analysis of work rate and heart rates of association football referees. *Br. J. Sports Med.* 27:153–156. 1993.
14. CHAMARI, K., S. AHMAIDI, C. FABER, AND J. MASSE-BIRON. Anaerobic and aerobic power output and the force-velocity relationship in endurance trained athletes: Effects of ageing. *Eur. J. Appl. Physiol.* 71:1230–1234. 1995.
15. D'OTTAVIO, S., AND C. CASTAGNA. Physiological load imposed on elite soccer referees during actual match play. *J. Sports Med. Phys. Fitness* 41:27–32. 2001.
16. D'OTTAVIO, S., AND C. CASTAGNA. Analysis of match activities in elite soccer referees during actual match play. *J. Strength Cond. Res.* 15:167–171. 2001.

17. D'OTTAVIO, S., AND C. CASTAGNA. Physiological aspects of soccer refereeing. In: *Science and Football IV*. W. Spinks, T. Reilly, and A. Murphy, eds. London: Routledge, 2002. pp. 144–150.
18. EISSMANN, H.J., AND M. D'HOOGHE. Sports medical examinations. In: *The 23rd Man: Sports Medical Advice for Football Referees*. H.J. Eissmann, ed. Leipzig: Gersöne-Druck, 1996. pp. 7–19.
19. FAY, L., B.R. LONDEREE, T.P. LaFontaine, AND M.R. VOLEK. Physiological parameters related to distance running performance in female athletes. *Med. Sci. Sports Exerc.* 21:319–324. 1989.
20. HARLEY, R.A., K. TOZER, AND J. DOUST. An analysis of movement patterns and physiological strain in relation to optimal positioning of Association Football referees. In: *Science and Football IV*. W. Spinks, T. Reilly, and A. Murphy, eds. London: Routledge, 2002. pp. 137–143.
21. HOFF, J., AND J. HELGERUD. Endurance and strength training for soccer players—Physiological considerations. *Sports Med.* 34:165–180. 2004.
22. JOHNSTON, L., AND L. McNAUGHTON. The physiological requirements of soccer refereeing. *Aust. J. Sci. Med. Sport.* 26(3/4):67–72. 1994.
23. KRUSTRUP, P., AND J. BANGSBO. Physiological demands of top-class soccer refereeing in relation to physical capacity: Effect of intense intermittent exercise training. *J. Sports Sci.* 19:881–891. 2001.
24. LUCIA, A., S.J. FLECK, R.W. GOTSHALL, AND T.J. KEARNEY. Validity and reliability of the Cosmed K2 instrument. *Int. J. Sports Med.* 14:380–386. 1993.
25. PIRNAY, F., AND J.M. CRIELAARD. Anaerobic power and ageing. In: *Frontiers in Sport Sciences, the European Perspective*. J.G.P. Marconnet, I. Margaritis, and F. Tessier, eds. Nice, France: University of Nice, 1996. pp. 302–303.
26. ROGERS, M.A., J.M. HAGBERG, W.H. MARTIN, AND N.A. ESHANI. Decline in VO₂max with ageing in masters athletes and sedentary men. *J. Appl. Physiol.* 68:2195–2199. 1990.
27. STAMFORD, B.A. Exercise and the elderly. *Exerc. Sports Sci. Rev.* 16:341–379. 1988.
28. WESTON, M., AND J. BREWER. A study of the physiological demands of soccer refereeing. *J. Sports Sci.* 20:59–60. 2002.
29. WILLIAMS, A.M., D. LEE, AND T. REILLY. *A Quantitative Analysis of Matches Played in the 1991–92 and 1997–98 Seasons*. London: The Football Association, 1999.

Address correspondence to Prof. Carlo Castagna, carlocastagna@teknosport.com.

4.2

STUDY V

Weston, M., Castagna, C., Impellizzeri, F., Rampinini, E. & Breivik, S. (2010)
Ageing and Physical Match Performance in English Premier League Soccer
Referees. *Journal of Science and Medicine in Sport*, 13, 96-100.

Original paper

Ageing and physical match performance in English Premier League soccer referees

Matt Weston^{a,*}, Carlo Castagna^b, Franco M. Impellizzeri^c,
Ermanno Rampinini^d, Simon Breivik^a

^a *The Professional Game Match Officials Limited (PGMOL), The FA Premier League, 30 Gloucester Place, London, UK*

^b *School of Sport and Exercise Sciences, Faculty of Medicine and Surgery, University of Roma Tor Vergata, Rome, Italy*

^c *Neuromuscular Research Laboratory Schulthess Klinik, Zurich, Switzerland*

^d *Human Performance Laboratory, S.S.MAPEI srl, Via Don Minzoni 34, Castellanza, Varese, Italy*

Received 28 January 2008; received in revised form 15 July 2008; accepted 30 July 2008

Abstract

Soccer referees are required to keep up with play at all times despite occupying an age bracket of on average 10–15 years older than their playing counterparts. Therefore, the aim of the present study was to examine the effect of age upon the physical match performances and match physiological loads of elite-level soccer referees. Match analysis data was collected (Prozone[®], Leeds, UK) from 22 professional soccer referees (age range 31–48 years) on FA Premier League matches over four consecutive seasons (778 observations). Physical match performance categories were total-distance covered (TD); high intensity running distance (speed $>5.5 \text{ m s}^{-1}$, HIR); sprint count ($>7.0 \text{ m s}^{-1}$, SC); top sprinting speed (TS); average distance from the ball (DB) and average distance from fouls (DF). Significant age effects were found for TD ($r = -0.52$, $p < 0.001$), HIR ($r = -0.53$, $p < 0.001$) and SC ($r = -0.53$, $p < 0.001$). No age effect was found for DB and DF ($p > 0.05$). Despite covering less TD, HIR and performing fewer sprints the older referees (43–48 years) were able to maintain an average distance from fouls that was comparable to that recorded by the young (31–36 years) referees. Therefore, the reduced physical match performances associated with increasing referee age did not appear to impact upon the older referees' ability to keep up with play. In light of these findings, refereeing governing bodies may wish to review their age-based retirement guidelines.

© 2008 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.

Keywords: Football; Match analysis; Positioning; Work rate; Experience

1. Introduction

The match activity profiles and physiological demands of soccer referees during competitive matches have been reported to be very similar^{1–3} and also related to those of the players.⁴ Consequently, the referee is required to keep up with play at all times despite occupying an age bracket on average 10–15 years older than their playing counterparts. The difference between the average age of players and referees may exist because experience is considered as a fundamental prerequisite to officiate at elite level.⁵ International refereeing governing bodies have an enforced retirement age of 45 years for referees. However, in some countries this enforced

retirement has been successfully challenged as it is considered to be against the European Employment Directive, which prohibits age discrimination in employment.

The scientific rationale behind an enforced retirement age limit for referees is not clear. One of the reasons may be the well documented age-related decline in physical fitness^{6–8} given the importance placed on a referees fitness levels.^{5,9} Any decline in fitness may impair a referee's ability to keep up with the play, yet in order to apply the laws of the game accurately a close proximity to fouls is considered important for fair judgement.^{1,3,10}

However, whilst significant age-related impairments in physical fitness have been reported in soccer referees^{11,12} the relationship between age and physical match performance has yet to be examined. Therefore, the aims of the present study were to (1) examine the effect of age upon the physi-

* Corresponding author.

E-mail address: mweston@pgmol.com (M. Weston).

cal match performances of soccer referees in relation to the distance from fouls and the ball over the course of four consecutive seasons; and (2) examine the effect of age upon the physiological load imposed upon soccer referees during competitive matches.

2. Methods

Data were collected from 22 professional English Football Association (FA) Premier League referees during FA Premier League matches throughout four seasons; 2003–2004, 2004–2005, 2005–2006 and 2006–2007, giving a total of 778 match observations. The referees' age was taken on the first day of each season. In line with previous studies^{11,12} the referees were assigned to three different age-group categories; young (31–36 years, $n=135$ match observations), intermediate (37–42 years, $n=308$ match observations) and older (43–48 years, $n=335$ match observations). However, as the retirement age on the English Premier League is 48 years as opposed to 45 years on the Italian and Spanish top leagues, the average age of the groups in the present study were higher than those previously reported.^{11,12} Written informed consent was received from all referees after verbal and written explanation of the experimental design and potential risks of the study. The local Institutional Review Board approved this study design.

In order to examine the effect of ageing upon the referees physical match performances and physiological stress the match activity, ratings of perceived exertion (RPE) and heart rate (HR) data from each of the referee's matches during each season were averaged to give a mean score for that particular year of age. Consequently the referees represented a total of 69 different age years (range 31–48 years) across the four soccer seasons considered. Therefore an average value for each of the variables of interest for each age year was obtained. However, due to retirement and also promotion, not all the referees involved contributed data on each of the four seasons (3.2 ± 1.1 seasons per referee). To further examine the possibility of age associated variations the match activity, RPE and HR data were assigned to the three referee age-group categories.

Each match was monitored using a computerised, semi-automatic video match analysis image recognition system (ProZone®, Leeds, England). The system has showed an almost perfect correlation with a timing gate measurement system at velocities ranging from 7.5 to 25 km h⁻¹ for 60 m runs ($r=0.99$), 50 m angled runs ($r=0.99$) and 20 m maximal sprints with a change of direction ($r=0.95$).¹³ The objective measures of match running performance selected for analysis were (1) total-distance covered (TD); (2) high intensity running distance (running speed >5.5 m s⁻¹ (>19.8 km h⁻¹), HIR); (3) sprint count (running speed >7.0 m s⁻¹ (>25.2 km h⁻¹), SC); (4) top sprinting speed (m s⁻¹, TS); (5) mean distance from the ball (m, DB); (6) mean distance from fouls (m, DF). Also, as the physical

match performances of English Premier League soccer referees have been demonstrated to be in part related to players' HIR⁴ the sum total of the players HIR distance (running speed >5.5 m s⁻¹ (>19.8 km h⁻¹), PHIR) during each match was also recorded as a measure of overall match intensity.

The referees recorded their RPE score, using Borg's CR10 scale,¹⁴ 30 min after the match had ended to obtain a global intensity rating for the entire match.^{15,16} Therefore, a particularly easy or difficult bout of exercise towards the end of the match would not dominate the referees' rating. This scale has been reported to correlate significantly with objective indices of exercise training such as HR and blood lactate.¹⁶ The referees were familiarised with the use of RPE scoring having used it over a 12-month period prior to the study. Match HR was recorded every 5 s and analysed via a short-range telemetry system (Polar S610 and Precision 3.0 Kemple, Finland, respectively).

Match RPE load was determined by multiplying the match duration (minutes) by the session RPE score.¹⁵ This method allows physiologists to quantify internal training load into a single term that balances exercise duration and intensity.¹⁷ Recent studies have demonstrated that the RPE load method is a valid tool for quantifying training load in team sports including soccer.^{15,18}

Match HR load was computed by multiplying the accumulated duration in each of five different HR zones by a multiplier for each zone ($<60\%$ HR_{max} = 1; 60–75% = 2; 76–85% = 3; 86–93% = 4; $>93\%$ = 5) and summing the results.¹⁹ The HR zones in this study were based on the individual physiological response to incremental exercise as described by Bourdon.²⁰ The referees' individual HR_{max} values were determined from the peak values reached in any of the 5 s periods observed during matches, training sessions or fitness tests.^{3,21}

Data are presented as the mean \pm S.D. The Shapiro-Wilk test was applied to the data in order to assess for a normal distribution of parametric data. Relationships between the referees' age and their physical match performances and physiological match responses were examined using a Pearson's product moment correlation. The following scale of magnitudes proposed by Hopkins²² was used to interpret the correlation coefficients: <0.1 , trivial; 0.1–0.3, small; 0.3–0.5, moderate; 0.5–0.7, large; 0.7–0.9, very large; >0.9 , nearly perfect. Levene test for equity of variances was computed with no significant differences being found. A one-way analysis of variance (ANOVA) was used to determine differences in physical match performance between the groups of referees. Analysis of covariance (ANCOVA) was performed assuming PHIR as covariate.⁴ Post hoc analyses were performed using Tukey's Unequal-N HSD test. Effect sizes (η^2) were also calculated and values of 0.01, 0.06 and above 0.15 were considered small, medium and large, respectively. Statistical significance was set at $p < 0.05$. All calculations were performed using the Statistica statistical analysis software package (Version 6.0).

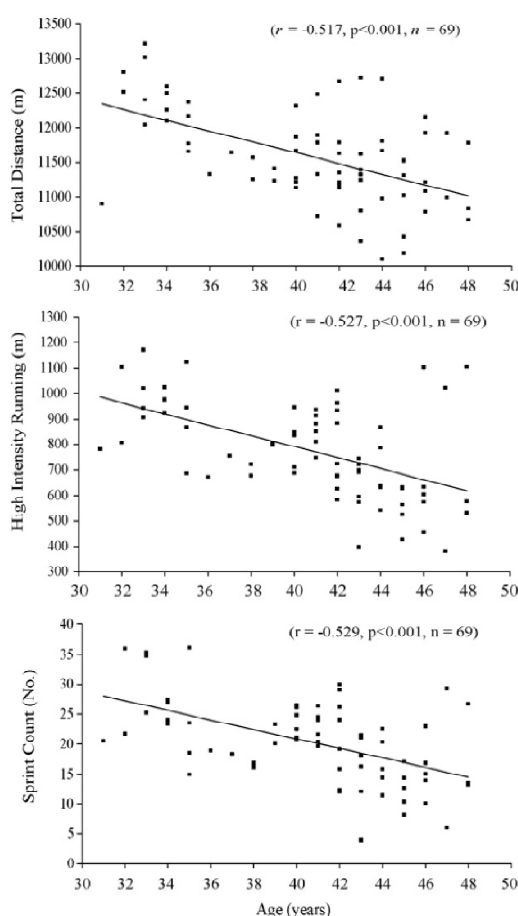


Fig. 1. Scatterplots of the significant correlations between referee age and physical match performance variables.

3. Results

Fig. 1 illustrates large negative correlations between referee age and the match variables of TD ($r = -0.52$ [CI 95%: $-0.67; -0.32$], $p < 0.001$, $n = 69$), HIR ($r = -0.53$ [CI 95%: $-0.68; -0.33$], $p < 0.001$, $n = 69$) and SC ($r = -0.53$ [CI 95%:

$-0.68; -0.33$], $p < 0.001$, $n = 69$). TS showed a small correlation with age ($r = -0.26$ [CI 95%: $-0.47; -0.02$], $p = 0.034$, $n = 69$). The match variables of DB and DF showed no relationship with age ($p > 0.05$) nor did the referees' match HR load ($p > 0.05$). Whereas, match RPE showed a moderate correlation with referee age ($r = 0.35$ [CI 95%: $0.12; 0.54$], $p < 0.01$, $n = 69$).

As can be seen from Table 1, ANOVA results demonstrated a significant age effect upon selected referee's physical match performance variables. Specifically, when split into three different age-group categories the young referees covered a greater TD when compared to the intermediate and older referees ($p < 0.05$), with HIR and SC both decreasing with increasing referee age group ($p < 0.05$). In terms of TS the older referees were significantly slower when compared to the young and intermediate groups ($p < 0.01$). No significant differences were found between the young and older group for DB and DF ($p > 0.05$).

With regards to overall match intensity there were no differences between the PHIR in the matches officiated by the three age groups ($p > 0.05$). Also, there were no between-group differences for match HR load ($F_{2, 709} = 0.22$, $p = 0.802$). Conversely, the referees match RPE load was significantly higher in the older group when compared to the young and intermediate ($p < 0.0001$) age groups. Effect size for all comparisons was small ($\eta^2 < 0.02$). ANCOVA results were not different from one-way ANOVA analysis.

4. Discussion

The results of the present study demonstrated an age-related decline in elite-level soccer referees physical match performances as TD, HIR and SC all decreased as referee age increased. Furthermore, the differences between young and older groups were greater for HIR and SC when compared to TD (-28.4 , -35% vs. -7.4% , respectively). This decline in physical match performance may be related to the age associated impairments in physical capacities already reported in soccer referees.^{11,12} Specifically, Casajus and

Table 1
Referees' match activity and physiological load during Premier League matches

	All $n = 778$	Young $n = 6$ (31–36 years) $n = 135$	Intermediate $n = 8$ (37–42 years) $n = 308$	Older $n = 8$ (43–48 years) $n = 335$
Match observations				
Total distance (m)	11534 ± 748	12209 ± 713 §	11490 ± 570 *	11302 ± 749
High intensity running (m)	773.1 ± 253.6	935.9 ± 250.7 §	814.8 ± 203.4 *	669.2 ± 251.6
Sprint count (no.)	20.1 ± 11.5	25.4 ± 12.6 §	21.7 ± 10.6 *	16.5 ± 10.6
Top speed (m s^{-1})	8.60 ± 0.52	8.69 ± 0.43 §	8.64 ± 0.50 *	8.53 ± 0.56
Players high intensity running (km)	19.5 ± 2.2	19.1 ± 2.3	19.6 ± 2.0	19.6 ± 2.4
Distance from ball (m)	19.3 ± 1.1	19.0 ± 1.2 §	19.5 ± 0.9 *	19.1 ± 1.2
Distance from fouls (m)	14.4 ± 1.6	13.9 ± 1.5 §	14.7 ± 1.6 *	14.3 ± 1.5
Match load				
RPE load (au)	645.6 ± 140.0	599.2 ± 132.3 §	623.1 ± 150.0 *	685.6 ± 122.7
HR load (au)	313.5 ± 35.2	314.8 ± 30.6	313.8 ± 42.7	312.5 ± 27.5

§ Significant difference between young and intermediate groups ($p < 0.05$). § Significant difference between young and older groups ($p < 0.05$). * Significant difference between intermediate and older groups ($p < 0.05$).

Castagna¹¹ and Castagna et al.¹² both reported a greater age-related decrement in anaerobic performance when compared to aerobic capacity. Our findings seem to be consistent with these results since the age-related decline in physical match performance was more pronounced amongst the predominantly anaerobic match variables of HIR and SC. Unfortunately, a limitation of the present study was that it was not possible to report on the referees' fitness levels as determined through regular fitness tests. Therefore, the link between the decline in physical fitness and physical match performance cannot be examined. However, irrespective of the reasons, the present study showed a moderate but significant age related decline in generic physical match performance.

No correlations were observed between age and the referees' DB and DF. A 2–3% lower DB and DF was found for older referees compared to intermediate but not younger referees. However, a difference of 0.4 m is of limited practical significance (small effect size: $\eta^2 < 0.02$) and may have been influenced by the large sample size of the study. Referees are required to keep up with play at all times and ensure optimal viewing positions.²³ In the present study the older referees reduced physical match performances did not affect their ability to keep up with play. As a result, it maybe the effectiveness of the referees match activities in ensuring the best possible viewing positions that is of most relevance. Expertise literature has demonstrated that experts are better than novices in using advance visual cues to guide their anticipatory responses.^{24,25} Therefore, our findings may be explained by the older, more experienced referees being better at anticipating and reading play and ultimately being more economical with their movements due to their many years of practice.

The relationship between DB and DF with viewing angle and ultimately correct decision making has yet to be examined and was beyond the scope of the present study. It maybe that the DF could be consistent between referees, yet the referees are viewing incidents from two entirely different angles, i.e., head-on and side-on. Consequently, this is an area that warrants further investigation given that no differences in DF were found despite wide ranging physical match performances.

The referees match HR load showed no relationship with age. Conversely, match RPE load demonstrated a moderate correlation with referee age ($r = 0.35$, $p < 0.01$, $n = 69$) and the older referees match RPE load was significantly higher when compared to the young and intermediate age groups despite covering less TD, HIR and performing fewer sprints. This may in part be explained by the fact that the experienced referees are more often assigned to officiate high profile matches, which may have led to an increase in the perceived match intensity, although no between-group differences were observed for PHIR. Alternatively, given the older referees in the present study perceived their match demands to be higher, despite covering less ground and at the same HR, the higher RPE might suggest greater fatigue.

The present studies findings would appear to suggest that the enforced retirement age of 45 years may not be justified by the inability of the referees to keep up with the play as the average age of the older group in this study was 44.8 years. Or, it may be that this enforced retirement age of 45 years has become dated with the increased fitness levels observed in referees over recent years after following more modern training regimes^{3,23} and also the advent of full-time, professional soccer referees.

5. Conclusion

The current findings demonstrated that reduced physical match performances with increasing referee age did not impact upon the older referees' ability to keep up with play. Such findings have a great deal of practical significance in a sport where the international governing body has an enforced retirement age of 45 years. It may be that the referees technical performances, fitness levels and physical match performances are all factors that should be taken into consideration when it comes to the assessment of whether a referee is 'fit for purpose'. Since the referees of the present study possibly experienced greater fatigue as indicated by the higher RPE load, future studies should examine if fatigue has an impact upon a referees decision-making process. Therefore, it is recommended that physical performances should be related to the decision-making process in an attempt to fully determine the impact of ageing upon match performance in soccer referees. Lastly, future studies should examine the relevance of distance from fouls in relation to viewing positions and ultimately correct decision making in soccer referees.

Practical implications

- Referees' physical match performance reduces with increasing age.
- Age does not appear to impact upon older referees' ability to keep up with play.
- Referees retirement should be undertaken considering ability to officiate and physical capacity.

Acknowledgements

We would like to acknowledge Martin Bland (Prozone[®]) for his help with the referees' match analysis data reduction. We sincerely thank Keith Hackett (FA Premier League Referees Manager) for his help and support in the project. The cooperation of the referees was of invaluable importance. No external financial support was required for this project.

References

1. Castagna C, Abt G, D'Ottavio S. Physiological aspects of soccer refereeing performance and training. *Sports Med* 2007;37(7):625–46.

2. Weston M, Bird S, Helsen W, Nevill A, Castagna C. The effect of match standard and referee experience on the objective and subjective match workload of English Premier League referees. *J Sci Med Sport* 2006;9(3):256–62.
3. Krstrup P, Bangsbo J. Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. *J Sports Sci* 2001;19:881–91.
4. Weston M, Castagna C, Impellizzeri F, Rampinini E, Abt G. Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates. *J Sci Med Sport* 2007;10(6):390–7.
5. Eissmann HJ. *The 23rd man: sport medical advice for football referees*. Leipzig (Germany): Gersöne-Druck; 1996.
6. Bruce RA. Exercise, aerobic capacity and ageing—another viewpoint. *Med Sci Sports Exerc* 1984;16:8–13.
7. Stamford BA. Exercise and the elderly. *Exerc Sport Sci Rev* 1988;16:341–79.
8. Chamari K, Ahmaidi S, Faber C, Masse Biron J, Prefaut C. Anaerobic and aerobic power output and the force–velocity relationship in endurance trained athletes: effects of ageing. *Eur J Appl Physiol* 1995;71:1230–4.
9. Mallo J, Navarro E, Garcia-Aranda JM, Gilis B, Helsen W. Activity profile of top-class association football referees in relation to performance in selected physical tests. *J Sports Sci* 2007;25(7):805–13.
10. Harley RA, Tozer K, Doust J. An analysis of movement patterns and physiological strain in relation to optimal positioning of Association Football referees. In: Spinks W, Reilly T, Murphy A, editors. *Science and football IV*. London: Routledge; 2002. p. 137–43.
11. Casajus JA, Castagna C. Aerobic fitness and field test performance in elite Spanish soccer referees of different ages. *J Sci Med Sport* 2007;10(6):382–9.
12. Castagna C, Abt G, D'Ottavio S, Weston M. Age-related effects on fitness performance in elite-level soccer referees. *J Strength Cond Res* 2005;19(4):785–90.
13. Di Salvo V, Collins A, McNeill B, Cardinale M. Validation of prozone: a new video-based performance analysis system. *Int J Perf Anal Sport* 2006;6(1):108–19.
14. Borg G, Hassmen P, Lagerstrom M. Perceived exertion related to heart rate and blood lactate during arm and leg exercise. *Eur J Appl Physiol Occup Physiol* 1987;56(6):679–85.
15. Foster C. Monitoring training in athletes with reference to overtraining syndrome. *Med Sci Sports Exerc* 1988;30(7):1164–8.
16. Foster C, Florhaug JA, Franklin J, Gottschall L, Hrovatin LA, Parker S, et al. A new approach to monitoring exercise training. *J Strength Cond Res* 2001;15(1):109–15.
17. Impellizzeri FM, Rampinini E, Marcora SM. Physiological assessment of aerobic training in soccer. *J Sports Sci* 2005;23(6):583–92.
18. Impellizzeri FM, Rampinini E, Coutts AJ, Sassi A, Marcora SM. Use of RPE-based training load in soccer. *Med Sci Sports Exerc* 2004;36(6):1042–7.
19. Edwards S. High performance training and racing. In: Edwards S, editor. *The heart rate monitor book*. Sacramento, CA: Feet Fleet Press; 1993. p. 113–23.
20. Bourdon P. Blood lactate transition thresholds: concepts and controversies. In: Gore CJ, editor. *Physiological tests for elite athletes*. USA: Human Kinetics; 2000. p. 50–66.
21. Helsen W, Bultynck JB. Physical and perceptual-cognitive demands of top-class refereeing in association football. *J Sports Sci* 2004;22:179–89.
22. Hopkins W. A new view of statistics. Available at: <<http://www.sportsci.org/resource/stats/index.html>> [accessed 22.09.07].
23. Weston M, Helsen W, MacMahon C, Kirkendall D. The impact of specific high intensity training sessions on Football Referees' fitness levels. *Am J Sport Med* 2004;32(1 Suppl.):54s–61s.
24. Williams AM, Burwitz L. Advance cue utilisation in soccer. In: Reilly T, Clarys J, J, Sibbe A, editors. *Science and football II*. London: E & FN Spon; 1993. p. 239–43.
25. Abernethy B. Anticipation in sport: a review. *Phys Ed Rev* 1987;10:5–16.

5.1**STUDY VI**

Weston, M, Castagna, C., Impellizzeri, F. & Helsen, W. (2009) Relationship between field tests and physical match performance in elite soccer referees.

Journal of Sports Sciences, 27 (11), 1177-1184.

Relationships among field-test measures and physical match performance in elite-standard soccer referees

MATTHEW WESTON¹, CARLO CASTAGNA², WERNER HELSEN³, & FRANCO IMPELLIZZERI⁴

¹PGMOL, The FA Premier League, London, UK, ²School of Sport and Exercise Sciences, University of Rome "Tor Vergata", Rome, Italy, ³Department of Biomedical Kinesiology, Katholieke Universiteit Leuven, Leuven, Belgium, and ⁴Neuromuscular Research Laboratory, Schulthess Clinic, Zurich, Switzerland

(Accepted 11 June 2009)

Abstract

The aim of this study was to assess the extent to which measures derived from the new FIFA referees' fitness tests can be used to monitor a referee's match-related physical capacity. Match-analysis data were collected (Prozone[®], Leeds, UK) from 17 soccer referees for 5.0 ($s = 1.7$) FA Premier League matches per referee during the first 4 months of the 2007–08 season. Physical match performance categories included total distance covered, high-intensity running distance (speed $> 5.5 \text{ m} \cdot \text{s}^{-1}$), and sprinting distance ($> 7.0 \text{ m} \cdot \text{s}^{-1}$). The two tests were a $6 \times 40\text{-m}$ sprint test and a 150-m interval test. Heart rate demand was correlated with total match distance covered ($r = -0.70$, $P = 0.002$) and high-intensity running ($r = -0.57$, $P = 0.018$) in the interval test. The fastest 40-m sprint was related to total distance covered ($r = -0.69$, $P = 0.002$), high-intensity running ($r = -0.76$, $P < 0.001$), and sprinting distance ($r = -0.75$, $P = 0.001$), while mean time for the 40-m sprints was related to total distance covered ($r = -0.70$, $P = 0.002$), high-intensity running ($r = -0.77$, $P < 0.001$), and sprinting distance ($r = -0.77$, $P < 0.001$). The referees who recorded the best interval-test heart rate demand and fastest 40-m time produced the best physical match performances. However, only the sprint test and in particular the fastest 40-m time had appropriate construct validity for the physical assessment of soccer referees.

Keywords: Soccer, match analysis, repeated-sprint ability, fitness, intermittent exercise

Introduction

National and international soccer referees' associations routinely assess the fitness of elite-standard officials because high fitness is considered to be necessary to cope with the physical stress imposed on referees during matches (Castagna, Abt, & D'Ottavio, 2002). Before 2005, referees had to pass a battery of field-based tests that included $2 \times 50\text{-m}$ and $2 \times 200\text{-m}$ sprints, followed by the 12-min Cooper test (Eissmann & D'Hooghe, 1996). However, research has illustrated that these tests are poor measures of referees' match-related physical capacity. Specifically, Castagna et al. (2002) reported that the 12-min run had only a low correlation with high-intensity and maximal speed running ($r = 0.51$ and $r = 0.32$, respectively). The 50-m and 200-m sprint times were not correlated with distances covered at high-intensity and maximal running speeds. Similar findings were reported by Mallo and colleagues (Mallo, Navarro, Garcia-Aranda, Gilis, & Helsen,

2007), who concluded that measures from the FIFA (Fédération Internationale de Football Association) fitness tests were not correlated with match activities and the evaluation of elite-standard referees should be more specific and related to activities performed during matches.

In an attempt to overcome this problem, FIFA introduced two new fitness tests: a repeated $6 \times 40\text{-m}$ sprint test and a high-intensity 150-m interval test (FIFA, 2008). Minimum requirements were set by FIFA so that should the referees achieve this standard, their fitness will be deemed sufficient to cope with match demands. However, the use of these tests as a measure of match-related fitness is unclear, as Mallo and colleagues (Mallo, Navarro, Garcia-Aranda, & Helsen, 2009) reported the 40-m mean and best sprint times to be poorly related to distance covered at high-speed running during matches ($r = -0.02$, $P = 0.96$ and $r = -0.13$, $P = 0.74$, respectively). Also, heart rate in the interval test was not correlated with total distance covered

($r = -0.07$, $P = 0.85$) or time spent running at high speed during matches ($r = -0.49$, $P = 0.18$).

These discrepancies could be attributable to technical issues associated with the new tests. For example, the determination of minimum standards requires empirical validation, and in the current test procedures there is no performance discrimination on the interval test and no precise resting time for the repeated-sprint test. If these issues were addressed – that is, the use of heart rate analysis to give a performance discriminator on the interval test and the standardization of the repeated-sprint recovery time – it would be possible to examine the extent to which measures derived from the tests could be used to assess referees' match-related physical capacity. Also, to increase the test demands FIFA decreased the recovery time on the interval test from 40 to 35 s (Mallo et al., 2009) and the validity of this altered protocol has yet to be determined.

Taking these issues into consideration, we hypothesized that: (1) physiological performance (i.e. heart rate response) on the interval test would be related to total distance covered and the distance covered at high-intensity running during matches; (2a) fastest 40-m sprint times would be related to the distance covered at maximal speeds during matches; and (2b) mean 40-m sprint would be related to the distance covered at high-intensity running during matches.

Methods

Participants

The participants were 17 English Football Association (FA) Premier League referees. Data were collected for FA Premier League matches during the first four months (August–November) of the 2007–08 English season. The mean age and body mass of the referees was 40.0 years ($s = 5.1$) and 82.8 kg ($s = 10.0$), respectively. Written informed consent was received from all referees after verbal and written explanations of the experimental design. The local Institutional Review Board approved the study design.

Experimental design

Physical match performances. The referees' physical match performances were examined for 85 matches (mean 5.0 matches per referee, $s = 1.7$). Each match was assessed using a computerized, semi-automatic video match-analysis image recognition system (Pro-Zone[®], Leeds, England). This method uses eight cameras positioned at roof height around football pitches to record positional data of every player and official, every 0.1 s during matches. The data captured are then systematically analysed using

proprietary software to provide comprehensive data on each individual, including distances covered and intensity of movement for each player and official. Recent findings have demonstrated that the Pro-Zone[®] match-analysis system provides valid and reliable analyses of movement patterns of footballers during match-play (Di Salvo, Collins, McNeill, & Cardinale, 2006; DiSalvo, Gregson, Atkinson, Tordoff, & Drust, 2009).

The objective measures of physical match performance selected for analysis were as follows: (1) total distance covered (m); (2) high-intensity running distance (running speed $> 5.5 \text{ m} \cdot \text{s}^{-1}$); (3) sprinting distance (running speed $> 7.0 \text{ m} \cdot \text{s}^{-1}$); (4) highest sprinting speed, calculated as the peak value recorded during 0.5-s samples ($\text{m} \cdot \text{s}^{-1}$); and (5) mean distance from fouls throughout the match (m), which has been suggested to be of technical importance in soccer refereeing, as being a considerable distance from fouls can lead to incorrect decisions (Krustrup & Bangsbo, 2001).

Fitness tests. Two fitness tests, which are the official FIFA referee fitness tests, were performed in the week before the start of the 2007–08 English Premier League season: a $6 \times 40\text{-m}$ sprint test and an interval test. The referees arrived at the test session having rested for a minimum of 48 h and having followed nutritional guidelines to ensure they were fully hydrated and energized. All referees had performed the tests on at least one occasion before the test session and had not reported any injuries throughout the period of pre-season training that preceded the fitness tests or throughout the period of match data collection.

(1) **$6 \times 40\text{-m}$ sprint test.** The FIFA test procedures state a maximum of 90 s recovery between each of the $6 \times 40\text{-m}$ linear sprints. However, the recovery time for the sprints in the present study was standardized at exactly 90 s. After a 25-min warm-up that consisted of low-intensity running, stretching, and striding, the referees performed the $6 \times 40\text{-m}$ sprint test, with 90 s of recovery between sprints. The referees' starting position was 1.5 m behind the starting line, in line with FIFA procedures, and the sprint times were recorded using photoelectric beams at 0 m and 40 m (Microgate, Italy). The referees were instructed to perform each sprint maximally and each sprint had to be completed in a maximum time of 6.2 s. The sprints were performed on a regular athletics track and the use of spikes was not allowed. If a referee failed one of the sprints, he was allowed an extra sprint. However, if a referee failed to run below 6.2 s on two occasions, he failed the test (FIFA, 2008).

The referees' fastest 40-m sprint time, mean $6 \times 40\text{-m}$ sprint time, and percent decrement in

40-m performance were used for analysis. The percent decrement in 40-m sprint performance was calculated following Spencer and colleagues (Spencer, Fitzsimmons, Dawson, Bishop, & Goodman, 2006).

After a recovery period of 8 min, the referees performed the interval test.

(2) *Interval test.* The referees' interval test alternated 150 m of running, which had to be completed in a maximum of 30 s, with 50 m of walking, which had to be completed in 35 s, around a 400-m athletics track (see Figure 1). During the interval runs as well as the recovery periods, an audio signal was played over a loudspeaker system warning the referees when 10 s (single bleep), 5 s (double bleep), and 0 s (whistle) were remaining (FIFA MP3 package).

The referees were required to complete a minimum of 10 laps to pass the test, which represented 20 high-intensity runs (4000 m in total, of which 3000 m consisted of high-intensity running). If a referee failed to complete a 150-m interval in 30 s on one occasion, they received a verbal caution. If they then failed to complete another 150-m interval in 30 s, they failed the test.

Heart rate demand. At present, FIFA's only selection criterion on the interval test is that the referees are required to complete 10 laps to. However, this does not permit any performance individualization. Therefore, although it is not a requirement of FIFA, individual heart rates during the test were recorded as a measure of physiological performance. Heart rates during the interval test were recorded using short-range radio telemetry (Polar S610, Kempele, Finland), with the data being recorded every 5 s. Data recording began at the start of the first lap and finished at the end of the 10th lap. After the test, the data were downloaded onto a computer for analysis (Polar Precision software, version 3.0, Kempele, Finland).

Interval test heart rate demand was computed by multiplying the accumulated duration in each of five different heart rate zones by a multiplier for each zone [$< 60\%$ maximal heart rate = 1; 60–75%

maximal heart rate = 2; 76–85% maximal heart rate = 3; 86–93% maximal heart rate = 4; $> 93\%$ maximal heart rate = 5] and summing the results (Edwards, 1993). The heart rate zones in this study were based on the individual physiological response to incremental exercise as described by Bourdon (2000). The referees' individual maximal heart rates were determined from the peak value recorded in any of the 5-s periods observed during matches, training sessions or fitness tests (Helsen & Bultynck, 2004). This method was used to obtain a true maximum value for the referees, as recent work has reported higher maximal heart rates during training and competition than laboratory fitness tests (Semin et al., 2008).

Statistical analyses

The data are presented as means and standard deviations (*s*). The assumption of normality was checked with the Shapiro-Wilk *W*-test. Relationships between the referees' fitness test performances and their physical match performances were examined using Pearson's product-moment correlation coefficient, with 95% confidence intervals also presented. The following scale of magnitudes proposed by Hopkins (2008) was used to interpret the correlation coefficients: < 0.1 = trivial; 0.10–0.29 = small; 0.30–0.49 = moderate; 0.50–0.69 = large; 0.70–0.90 = very large; > 0.90 = nearly perfect. Levene's test confirmed equity of variances. A one-way repeated-measures analysis of variance (ANOVA) was performed on the groups' 6 \times 40-m sprints. The median split technique was used to determine between-group differences. This procedure produces two groups based on performance on each of the fitness tests. The referees were assigned to an upper or lower group depending on whether their score was below or above the median value of the entire sample, respectively. The median value was assigned to the upper group. Upper and lower groups were assigned using the median value for interval test heart rate demand, fastest 40-m sprint time, mean 6 \times 40-m sprint time, and percent decrement in 40-m performance. After this procedure, the physical match performance variables (total distance, high-intensity running, sprinting distance, highest sprinting speed, and distance from fouls) of the upper group were compared with those of the lower group using unpaired *t*-tests. Effect sizes (*d*) for the between-group differences were also determined, with values of 0.20–0.49, 0.50–0.80, and > 0.80 representing a small, moderate, and large difference, respectively (Vincent, 1995). Statistical significance was set at $P < 0.05$. All calculations were performed using the Statistica statistical analysis software package (Version 6.0, Statsoft Inc., Tulsa, OK, USA).

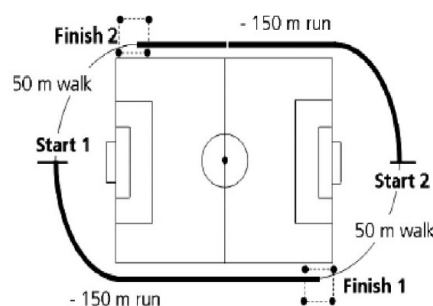


Figure 1. Schematic illustration of the FIFA referees' interval test.

Results

Fitness tests

All referees completed exactly 20 interval runs (10 laps) on the interval test (total test time of 21 min and 40 s). The heart rate demand for the 10 laps was 82.3 au ($s=9.1$, range = 67–98). The fastest 40-m sprint time was 5.59 s ($s=0.21$, range = 5.25–5.87) and mean 6×40 -m sprint time was 5.71 s ($s=0.19$, range = 5.37–5.95). In terms of the percent decrement in 40-m performance, the decrease was 2.0% ($s=0.94\%$, range = 0.7–4.3). Each referee successfully achieved the FIFA standard on all 6×40 -m sprints.

The referees demonstrated no change in sprint time across the six repeated-sprints ($F_{5,101}=0.32$, $P=0.90$). There was a nearly perfect correlation between fastest 40-m sprint time and mean 6×40 -m sprint time ($r=0.98$, $P<0.001$). Correlations were also observed between interval test heart rate demand and each of fastest 40-m sprint time ($r=0.67$, $P=0.003$), mean 6×40 -m sprint time ($r=0.58$, $P=0.014$), and percent decrement in 40-m performance ($r=-0.70$, $P=0.002$).

Physical match performances

The total distance covered by referees was 11,478 m ($s=580$, range = 10,556–12,640). The distance

covered at different running speeds was 753 m ($s=231$, range = 497–1177) for high-intensity running and 160 m ($s=78$, range = 66–353) for sprinting. The highest sprinting speed was $8.83 \text{ m} \cdot \text{s}^{-1}$ ($s=0.02$, range = 8.30–9.18) and the average distance from fouls was 14.3 m ($s=1.1$, range = 12.6–16.6).

Relationship between fitness tests and the referees' physical match performances

The correlation coefficients between fitness test results and physical match performance variables are presented in Figure 2 and Table I. Very large correlations were observed between interval test heart rate demand and total match distance, fastest 40-m sprint time and, high-intensity running and sprinting distance, as well as between mean 6×40 -m sprint time and each of total distance, high-intensity running, and sprinting distance. A large correlation was observed for interval test heart rate demand and high-intensity running, as well as for fastest 40-m sprint time and total distance covered. Large correlations were also observed between fastest 40-m sprint time and mean 6×40 -m sprint time with highest sprinting speed. The referees' distance from fouls was not related to any of the fitness test performance variables.

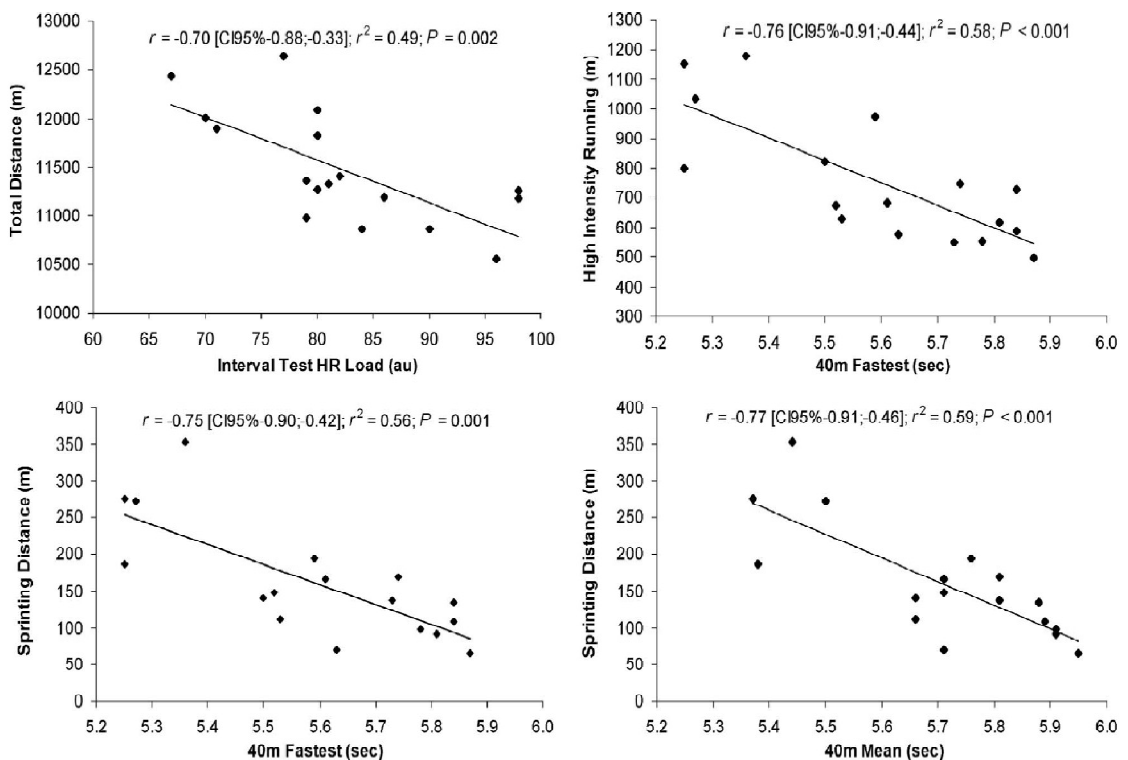


Figure 2. Scatterplots of the very large correlations between measures derived from the FIFA referees' interval test and 6×40 -m sprint test with match physical performance variables ($n=17$).

Table II illustrates that referees in the upper group for interval-test heart rate demand and fastest 40-m sprint time covered greater total distance, performed more high-intensity running and sprinting, and also had a higher sprinting speed than the referees in the lower group. Effect sizes for these comparisons were at least large. However, for mean 6×40 -m sprint time, the faster referees covered greater total distance but there were no between-group differences for all other variables, even though the effect sizes were large for the variables and moderate for highest sprinting speed. There were no between-group differences for percent decrement in 40-m performance.

Discussion

In the present study, we found that the new FIFA referees' fitness tests did demonstrate some measure

of a referee's match-related fitness. Specifically, heart rates in the interval test demonstrated a large correlation with total distance covered and match high-intensity running. These results are in contrast to those of Mallo et al. (2009), which could reflect the differences in recovery interval duration, as from 2006 onwards FIFA reduced the time from 40 to 35 s and the data reported by Mallo et al. (2009) were collected in 2005. There were very large correlations between performance on the 40-m sprint test and both match high-intensity running and sprinting despite the fact that Mallo et al. (2009) reported that the best and mean 40-m times were poorly correlated with the distance covered at high-speed running during matches. It could be that the faster high-intensity speed threshold used in the present study improved the sensitivity with which the relationship between test performance and match

Table I. Correlation coefficients between measures derived from the FIFA referees' interval test and 6×40 -m sprint test with match physical performance variables ($n = 17$).

	Total distance (m)	High-intensity running (m)	Sprinting distance (m)	Highest speed ($\text{m} \cdot \text{s}^{-1}$)
Interval test heart rate	-0.70**	-0.57*	-0.47	-0.27
40m _{fastest}	-0.69**	-0.76***	-0.75**	-0.52*
40m _{mean}	-0.70**	-0.77***	-0.77***	-0.50*
40m _{decrement}	-0.41	-0.41	-0.36	-0.34

Note: 40m_{fastest} = fastest 40-m sprint time; 40m_{mean} = mean 6×40 -m sprint time; 40m_{decrement} = percent decrement in performance.

*** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

Table II. Distance covered (mean \pm s) for different match activity classifications in two groups of elite-standard soccer referees (upper, $n = 9$; lower, $n = 8$) according to the median value during the interval test and 6×40 -m sprint-test (effect sizes, d , are also presented).

	Total distance (m)	High-intensity running (m)	Sprinting distance (m)	Highest speed ($\text{m} \cdot \text{s}^{-1}$)	Distance from fouls (m)
Interval-test heart rate					
Upper	11,833 \pm 546	883 \pm 209	205 \pm 78	8.93 \pm 0.12	14.0 \pm 0.8
Lower	11,079 \pm 290	607 \pm 88	109 \pm 36	8.70 \pm 0.24	14.5 \pm 1.3
Difference	753	276	96	0.23	-0.44
P -value	0.003	0.004	0.007	0.031	0.422
d	1.30	1.29	1.23	1.09	0.42
40m _{fastest}					
Upper	11,833 \pm 546	883 \pm 209	205 \pm 78	8.93 \pm 0.12	14.0 \pm 0.8
Lower	11,079 \pm 290	607 \pm 88	109 \pm 36	8.70 \pm 0.24	14.5 \pm 1.3
Difference	753	276	96	0.23	-0.44
P -value	0.003	0.004	0.007	0.031	0.422
d	1.30	1.29	1.23	1.09	-0.42
40m _{mean}					
P -value	11,741 \pm 580	838 \pm 229	191 \pm 91	8.87 \pm 0.21	14.3 \pm 1.1
d	11,182 \pm 440	657 \pm 155	125 \pm 42	8.77 \pm 0.22	14.1 \pm 1.0
P -value	559	181	66	0.11	0.19
d	0.040	0.074	0.073	0.332	0.716
P -value	0.96	0.85	0.85	0.49	0.18
40m _{decrement}					
Upper	11,308 \pm 568	684 \pm 202	144 \pm 87	8.73 \pm 0.23	14.3 \pm 1.3
Lower	11,669 \pm 566	830 \pm 210	178 \pm 67	8.93 \pm 0.15	14.2 \pm 0.8
Difference	-361	-145	-34	-0.20	0.06
P value	0.210	0.169	0.380	0.055	0.913
d	0.62	0.68	0.44	0.91	0.05

Note: 40m_{fastest} = fastest 40-m sprint time; 40m_{mean} = mean 6×40 -m sprint time; 40m_{decrement} = percent decrement in 40-m sprint performance.

performance could be assessed. Also, Mallo et al. (2009) did not have a separate classification for sprinting. Furthermore, differences could also be attributed to the low statistical power of the studies, as Mallo et al. (2009) reported simultaneous recordings of match analysis and physical capacity for only nine matches.

The referees' interval test heart rates correlated with the total distance covered by the referees ($r = -0.70$, $P = 0.002$). This finding is in line with the relationship between match distance and maximal performance on the 12-min run (Castagna et al., 2002), the Yo-Yo intermittent recovery test (Krustrup & Bangsbo, 2001), as well as maximum oxygen uptake (Castagna & D'Ottavio, 2001). However, since standing, walking, and jogging account for more than 75% of a referee's match activity (Krustrup & Bangsbo, 2001), it is the amount of high-intensity exercise that best indicates the demands of the game and the development of fatigue (Mohr, Krustrup, & Bangsbo, 2005). Indeed, a large correlation between interval test heart rate demand and match high-intensity running was observed ($r = -0.57$, $P = 0.018$).

In addition to the large correlation between interval test heart rate demand and match high-intensity running, Table II illustrates that referees in the upper group for interval test heart rate demand had better physical match performances than those in the lower group. However, a field test cannot be considered sport specific until a direct association between the most relevant aspects of physical match performance and the field test investigated has been determined (Castagna, Impellizzeri, Rampinini, D'Ottavio, & Manzi, 2008). This is evidenced by a common variance of only 33% between the two variables and referees' wide-ranging physical match performances (high-intensity range: 563–1530 m) despite the fact that all of them achieved FIFA's minimum standard of 10 laps. However, as referees' physical match performances are partly related to the intensity of the game itself (Weston, Castagna, Impellizzeri, Rampinini, & Abt, 2007), it is acknowledged that inter-match differences in intensity could account for part of this variation. Also, the interval test could have demonstrated a stronger relationship with physical match performance had the referees run to exhaustion. However, given that the test is not incremental, this would involve a high volume of running by referees in the best physical condition. Individual heart rate demand in the interval test provided an individual performance discriminator and, given our results, referees' heart rate response should be part of FIFA's selection criteria provided an accurate maximal value can be obtained for all referees before the test. Ultimately, further validation and possible modifications on the selection criteria

for the interval test are required if it is to be considered a truly valid measure of physical match performance in elite-standard soccer referees.

The Yo-Yo intermittent recovery test is considered a specific soccer refereeing test, since it has been reported to possess construct validity (Castagna, Abt, & D'Ottavio, 2005) and training sensitivity (Krustrup & Bangsbo, 2001; Weston, Helsen, MacMahon, & Kirkendall, 2004) in elite-standard soccer referees. In the interest of worldwide uniformity, FIFA require all fitness tests to be performed on a standard athletics track with limited specialized testing equipment. However, the Yo-Yo test is also performed on a consistent surface and requires the same specialist equipment (i.e. a pacing signal and audio speakers). Therefore, the use of the Yo-Yo test should be taken into consideration, although the setting of a FIFA minimum standard would still require empirical validation.

It was hypothesized that the sprint test scores would be related to distance covered at high-intensity running and sprinting during matches. Very large correlations were found between the referees' performance on the sprint test and these distances. Performance on the 40-m sprint test correlated with total distance covered and there were also large correlations between interval test heart rate demand and both fastest 40-m sprint time and mean 6×40 -m sprint time. These findings demonstrate a clear relationship between performance on the two tests, because those referees who performed better on the interval test tended to produce the fastest 40-m sprint times. This could reflect a predominance of repeated-sprint training in referees' training schedules, since Bravo et al. (2008) reported that such training is an effective strategy for improving both aerobic and anaerobic soccer-specific fitness.

These findings are also apparent in Table II, which illustrates that those referees in the upper group for fastest 40-m sprint time had better physical match performances than the lower group. Notably, despite a common variance of 95% between these two measures, those referees with the faster mean 6×40 -m sprint time performed better only on total distance covered. There were no between-group statistical differences for any other match variables despite large effect sizes and P -values approaching conventional standards of statistical significance. This suggests the occurrence of type II errors, low statistical power, and improved metrics provided by effect sizes.

Through the use of the median split technique, Rampinini et al. (2007) reported no differences in fastest 40-m sprint time between upper and lower groups for high-intensity running and sprinting. However, they did report between-group differences for the same match variables mean 40-m sprint time.

Although there were very large correlations between mean 6×40 -m sprint time and these match variables, we found no between-group differences based on mean 6×40 -m sprint time for high-intensity running and sprinting. Rampinini et al. (2007) suggested that their findings justified the use of a repeated-sprint test rather than a single sprint test to assess soccer players. This is likely because the physiological attributes required to perform the repeated-sprint test are similar to those required to maintain high-speed and sprinting performance throughout soccer matches (Rampinini et al., 2007). The difference between these findings and those reported in the present study could be attributable to the different techniques used. Rampinini et al. (2007) used a single sprint immediately before the repeated-sprint test and this trial provided the criterion score for the subsequent test. This procedure is not included in the FIFA referees' sprint test and despite being instructed to perform each sprint maximally, it is difficult to assess whether the referees did so, especially as some referees can achieve the selection criterion with sub-maximal efforts. This questions the rationale behind the minimum standards and therefore the test protocol might not provide an accurate reflection of referees' repeated-sprint ability.

The inclusion of the percent decrement in 40-m sprint performance as a test measure is questionable because in the present study there was no relationship between this measure and any of the physical match performance variables. Moreover, the reliability of this measure is low (Impellizzeri et al., 2008; Spencer et al., 2006).

The recovery time between sprints in the present study was 90 s, which is longer than that reported in previous studies that addressed repeated-sprint ability (Rampinini et al., 2007; Spencer et al., 2006). Consequently, the group demonstrated no change in sprint time across the six sprints and subsequently the group percent decrement in 40-m performance was less than in previous studies of repeated-sprint performance (Rampinini et al., 2007; Spencer et al., 2006). The absence of any performance decrement with 90 s recovery is consistent with the results of Balsom and colleagues (Balsom, Seger, Sjodin, & Ekblom, 1992), who reported that 40-m sprint time did not decrease until the 12th sprint with 60 s recovery. Furthermore, a recovery time of 90 s equates to an exercise-to-rest ratio of 1:14, which is higher than the critical 1:10 previously reported for soccer referees (Abt, Castagna, Belardinelli, & McCarthy, 2004). Therefore, to accurately measure repeated-sprint ability in soccer referees, the recovery time on the FIFA sprint test should be decreased and the test protocol harmonized with previously validated tests (Balsom et al., 1992;

Rampinini et al., 2007; Spencer et al., 2006). The assessment of fatigue is one of the key outcomes from a multiple-sprint test (Glaister, Howatson, Pattison, & McInnes, 2008), and if the current repeated-sprint test does not provide this measure, then its inclusion in the test procedure is questionable given the strength of the relationship between fastest 40-m sprint time and both match high-intensity running and sprinting.

In summary, given the relevance of high-intensity running for a referee's physical performance, only the sprint test and in particular the fastest 40-m sprint time had appropriate construct validity for the physical assessment of soccer referees. As national and international soccer referees' associations use the FIFA tests as selection criteria for competitions, modifications of the test protocols should be considered to increase their construct and logical validity.

Acknowledgements

We would like to acknowledge Martin Bland (Prozone®) for his help with the referees' match-analysis data reduction. We sincerely thank Keith Hackett (FA Premier League Referees Manager) for his help and support with the project. The cooperation of the referees was invaluable.

References

- Abt, G., Castagna, C., Belardinelli, R., & McCarthy, J. (2004). The effect of sprint duration on repeated sprint exercise with a constant work to rest ratio. *Journal of Sports Sciences*, 22, 521.
- Balsom, P., Seger, J. Y., Sjodin, B., & Ekblom, B. (1992). Maximal-intensity intermittent exercise: Effect of recovery duration. *International Journal of Sports Medicine*, 13, 528–533.
- Bourdon, P. (2000). Blood lactate transition thresholds: Concepts and controversies. In C. J. Gore (Ed.), *Physiological tests for elite athletes* (pp. 50–66). Champaign, IL: Human Kinetics.
- Bravo, D. F., Impellizzeri, F. M., Rampinini, E., Castagna, C., Bishop, D., & Wisløff, U. (2008). Sprint vs. interval training in football. *International Journal of Sports Medicine*, 29, 668–674.
- Castagna, C., Abt, G., & D'Ottavio, S. (2002). Relationship between fitness tests and match performance in elite soccer referees. *Journal of Strength and Conditioning Research*, 16, 231–235.
- Castagna, C., Abt, G., & D'Ottavio, S. (2005). Competitive-level differences in Yo-Yo intermittent recovery and 12 min run test performance in soccer referees. *Journal of Strength and Conditioning Research*, 19, 805–809.
- Castagna, C., & D'Ottavio, S. (2001). Effect of maximal aerobic power on match performance in elite soccer referees. *Journal of Strength and Conditioning Research*, 15, 420–425.
- Castagna, C., Impellizzeri, F. M., Rampinini, E., D'Ottavio, S., & Manzi, V. (2008). The Yo-Yo intermittent recovery test in basketball players. *Journal of Science and Medicine in Sport*, 11, 202–208.
- Di Salvo, V., Collins, A., McNeill, B., & Cardinale, M. (2006). Validation of Prozone: A new video-based performance analysis system. *International Journal of Performance Analysis in Sport*, 6, 108–119.

- Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P., & Drust, B. (2009). Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine*, 30, 205–212.
- Edwards, S. (1993). High performance training and racing. In S. Edwards (Ed.), *The heart rate monitor book* (pp. 113–123). Sacramento, CA: Feet Fleer Press.
- Eissmann, H. J., & D'Hooghe, M. D. (1996). Sports medical examinations. In H. J. Eissmann (Ed.), *The 23rd man: Sport medical advice for football referees* (pp. 7–19). Leipzig: Gersöne-Druck.
- Fédération Internationale de Football Association (2008). *Refereeing 2008: International lists* (pp. 246–249). Zurich: FIFA.
- Glaister, M., Howatson, G., Pattison, J. R., & McInnes, G. (2008). The reliability and validity of fatigue measures during multiple-sprint work: An issue revisited. *Journal of Strength and Conditioning Research*, 22, 1597–1601.
- Helsen, W., & Bultynck, J. B. (2004). Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, 22, 179–189.
- Hopkins, W. (2008). *A new view of statistics*. Retrieved April 22, 2008, from <http://www.sportsci.org/resource/stats/index.html>
- Impellizzeri, F. M., Rampinini, E., Bishop, D., Castagna, C., Ferrari Bravo, D., Tibaudi, A. et al. (2008). Validity of a repeated sprint test for football. *International Journal of Sports Medicine*, 29, 899–905.
- Krustrup, P., & Bangsbo, J. (2001). Physiological demands of top-class soccer refereeing in relation to physical capacity: Effect of intense intermittent exercise training. *Journal of Sports Sciences*, 19, 881–891.
- Mallo, J., Navarro, E., Garcia-Aranda, J. M., Gilis, B., & Helsen, W. (2007). Activity profile of top-class association football referees in relation to performance in selected physical tests. *Journal of Sports Sciences*, 25, 805–813.
- Mallo, J., Navarro, E., Garcia-Aranda, J. M., & Helsen, W. (2009). Activity profile of top-class association football referees in relation to fitness-test performance and match standard. *Journal of Sports Sciences*, 27, 9–17.
- Mohr, M., Krustrup, P., & Bangsbo, J. (2005). Fatigue in soccer: A brief review. *Journal of Sports Sciences*, 23, 593–599.
- Rampinini, E., Bishop, D., Marcora, S. M., Ferrari Bravo, D., Sassi, R., & Impellizzeri, F. M. (2007). Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. *International Journal of Sports Medicine*, 28, 228–235.
- Semin, K., Stahlnecker, A. C., Heelan, K., Brown, G. A., Shaw, B. S., & Shaw, I. (2008). Discrepancy between training, competition and laboratory measures of maximum heart rate in NCAA division 2 distance runners. *Journal of Sports Science and Medicine*, 7, 455–460.
- Spencer, M., Fitzsimmons, M., Dawson, B., Bishop, D., & Goodman, C. (2006). Reliability of a repeated-sprint test for field hockey. *Journal of Science and Medicine in Sport*, 9, 181–184.
- Vincent, W. J. (1995). *Statistics in kinesiology*. Champaign, IL: Human Kinetics.
- Weston, M., Castagna, C., Impellizzeri, F. M., Rampinini, E., & Abt, G. (2007). Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates. *Journal of Science and Medicine in Sport*, 10, 390–397.
- Weston, M., Helsen, W., MacMahon, C., & Kirkendall, D. (2004). The impact of specific high-intensity training sessions on football referees' fitness levels. *American Journal of Sport Medicine*, 32, S54–S61.

6.0 SUMMARY AND IMPLICATIONS FOR FUTURE RESEARCH

In this section the manuscripts that make up the PhD submission have been divided into four areas: 1) training performance, 2) match demands, 3) ageing and performance and, 4) fitness and match performance.

The data collection for this thesis commenced in late 2001. However, prior to this an initial study was designed to examine the physiological demands of soccer refereeing in England at the elite-level. Match heart rate and laboratory VO_2 data were collected during the 2000/2001 English soccer season. These data were presented at the 2001 British Association of Sport and Exercise Sciences (BASES)¹ annual conference. Following laboratory assessments the referees' mean $\text{VO}_{2\text{max}}$ was $50.9 \pm 5.7 \text{ ml.kg}^{-1}.\text{min}^{-1}$ with a HR_{max} of $178 \pm 10 \text{ b.min}^{-1}$. The mean heart rate response for 24 matches was $153 \pm 6.7 \text{ b.min}^{-1}$, the equivalent to $86.9 \pm 3.8\%$ of HR_{max} . The study demonstrated that, through the use of the laboratory measured heart rate- VO_2 relationship, the referees aerobic system was heavily taxed during game play as the estimated mean VO_2 for matches was 80.5% of $\text{VO}_{2\text{max}}$. Estimated energy expenditure during games was calculated using the assumption that 1 litre of oxygen consumed is equal to 5 kcal. A total of $1702 \pm 113 \text{ kcal}$ were expended during an average 97-min (inc. 7 min injury time) match, with $846 \pm 107 \text{ kcal}$ and $855 \pm 119 \text{ kcal}$ of energy expended in the first and second half, respectively. From this initial study it was clear that soccer refereeing involves a high physiological and energetic demand and therefore referees should possess a high level of aerobic fitness in order to

¹ Weston, M. & Brewer, J. (2002) An investigation into the physiological demands of soccer refereeing. *Journal of Sports Sciences*, 20, 59-60.

cope with the match demands. Consequently training programmes should be designed with this in mind.

6.1 TRAINING PERFORMANCE IN SOCCER REFEREES

STUDY I

Weston, M., Helsen, W., MacMahon, C. & Kirkendall, D. (2004) The impact of specific high intensity training sessions on football referees' fitness levels. *American Journal of Sports Medicine*, 32, S54-S61.

The purpose of this study was to assess the trainability of soccer referees using a training regime of predominantly high-intensity training activities. It was found that a structured training programme of twice weekly intensive intermittent training sessions significantly improved the fitness levels of soccer referees, as determined by a discontinuous incremental, field-based fitness test (YoYo Intermittent Recovery test level 1; YYIR1). A unique aspect of this study was that the previous study examining the trainability of soccer referees was short-term in its nature (12 weeks) (Krustrup & Bangsbo, 2001), whereas the present study spanned a period of 18 months. Also, training intensity was reported only as the prescribed intensity in the Krustrup & Bangsbo study (high-intensity running at 90% HR_{max}); whereas the present study progressed the work of Krustrup & Bangsbo (2001) by reporting in-session heart rates. Therefore, for the first time the medium-term effects of high-intensity training upon referees' fitness levels were reported with the results demonstrating visibly the impact of such training regimes.

When training referees it is important not to lose sight of the main goal; improving match performance. It is of little use improving fitness levels if it does not produce subsequent improvements in match performance. Unfortunately a weakness of our research design was that it was not possible to report on the referees' physical match performances before and after the training intervention, as per the work of Krstrup & Bangsbo (2001). Also it was not possible to recruit an age-matched control group, which will have decreased the internal validity of the study's findings and the inconsistent scheduling of the tests was another limitation; a common problem when working with elite sports performers (Atkinson & Nevill, 2001). However, we were confident that the improvements in the referees' YYIR1 performance levels will have had a positive impact upon the field of play given that the test is highly correlated with a referees' physical match performance (Krstrup & Bangsbo, 2001). Also the training session intensity was consistent with the intensity reported for soccer players and referees who improved physical match performance following a short-term, high-intensity training intervention (Helgerud, Engen, Wisloff & Hoff, 2001; Krstrup & Bangsbo, 2001). Therefore, such an improvement in fitness will have helped improve the referees' match performance through an increased ability to perform high-intensity running, thus enabling them to better keep up with play and to obtain optimal viewing positions in relation to critical match incidents.

Whilst this applied research study suffered from the lack of control that would be associated with a more basic piece of research, it did address a real-world issue, that is, the impact of up-to-date training methods on referees' fitness levels. The study also provided results of direct relevance to the practitioner; results that

helped to inform future practice with regards to training plan design and referee education. For example, the actual training session heart rates collected during this study demonstrated to the referees for the first time that it was possible to attain the same training stimulus during intensive intermittent field-based training sessions, when compared to the more ‘traditional’ track-based and road running exercises. This occurred despite the field-based sessions being shorter in duration and in a sport where only a small number are fortunate to be full-time, time/intensity effective training sessions are crucial if a referee is to be able to maximise their training output. Also, the improvement in YYIR1 performance observed by both groups of referees provided clear evidence as to the impact upon fitness levels of more up-to-date training methods and these data have consequently helped to educate referees at all levels on the effectiveness of high-intensity intermittent field-based training sessions.

6.2 MATCH DEMANDS OF ELITE-LEVEL SOCCER REFEREEING

STUDY II

Weston, M., Bird, S., Helsen, W., Nevill, A. & Castagna, C. (2006) The effect of match standard and referee experience upon the objective and subjective match workload of English Premier League referees. *Journal of Science and Medicine in Sport*, 9, 256-262.

Previous research examining the heart rate responses of referees during professional English soccer matches had expressed the heart rate recordings as a % of HR_{max} (Catterall, Reilly, Atkinson & Coldwells, 1993). However, in the

Catterall *et al.* study HR_{max} was estimated using the 220-age formula, a formula that overemphasises the age-related reduction in heart rate (Gellish *et al.*, 2007). Therefore, a strength of the present study was that the referees' match heart rates were expressed as a % of a laboratory determined maximal heart rate (HR_{max}). Also, up to this point no study had reported referees' ratings of perceived exertion (RPE), a valid measure of exercise intensity, on competitive matches.

This study reinforced earlier work in that soccer refereeing places a high physical load on referees, with an average Premier League match heart rate of $83.6\% \pm 0.8 HR_{max}$ and an average RPE score of 7.8 ± 0.8 . It was also found that match intensity was related to the standard of competition officiated as the same referees recorded lower heart rate and RPE scores when officiating on Football League matches ($81.5 \pm 2.2\%HR_{max}$, Effect size = 0.8; RPE 6.9 ± 0.8 , Effect size = 1.0).

The results presented in Study II are in contrast to those reported by Catterall *et al.* (1993) who, using data collected from Premier, Football and non-Football League matches, reported that the mean heart rate was unaffected by the category of competition, although higher peak heart rates were recorded during games in the top division. However, heart rate data were only collected from one match per referee. When comparing the mean match heart rates of top-class and high-standard referees Krstrup & Bangsbo (2001) reported no significant differences (164 ± 2 vs. 161 ± 3 b.min⁻¹).

Therefore, an important aspect of the research design in Study II was that match data were collected from the same referees at the two different standards of competition, with data collected on each Football League match taking place within 14 days of a Premier League match. Unfortunately, it was not possible to complete the research design by collecting data from Football League referees on Premier League matches. No experience effect on Premier League match heart rate and RPE scores was reported in the present study, although the effect size for the heart rate comparisons between low and high experience was large for both Premier League (0.8) and Football League (0.7). Effect size for RPE comparisons was small at both standards of competition (0.1 and 0.2, respectively).

The lower heart rates recorded by the low-experience referees may have reflected greater fitness levels, given that a faster heart rate recovery is concomitant with improved fitness levels. Therefore, it is important that the referees' match intensity scores are placed in context with the referees' fitness levels to obtain an accurate reflection of whether or not the referees met the physical demands of their match; although referees, like players, may not utilise full their physical capacity during matches.

These results of Study II have practical implications for a referee's fitness preparation and evaluation, in that referee and football governing bodies should ensure that appropriate levels of match fitness are developed to enable referees to cope with the increase in match demands that accompanies an elevation in competition standard. The Professional Game Match Officials Limited

(PGMOL), who oversee the training and appointment of all the match officials involved in the professional English game, now ensure that such procedures are adhered to when they promote their referees to a higher standard of competition. Once nominated for promotion the referees undergo a fitness assessment and have to attain the minimum standard of the level that they have been promoted to – if this is not achieved then the promotion is rescinded. Also, the referees' training heart rates and ratings of perceived exertion are analysed in detail by sports scientists to ensure a training programme of sufficient intensity and volume that will serve to promote the necessary fitness adaptations.

It was also concluded from the study that given the strength of the relationship between match RPE scores and mean match heart rates, match heart rate may be considered a valid indicator of global match intensity. However, this finding may have been overemphasised given a common variance of only 24% between the two variables. Recent literature has demonstrated a much stronger correlation between heart rate and RPE during training sessions when compared to competitive exercise. For example, Coutts, Rampinini, Marcora, Castagna & Impellizzeri (2007) reported that 43.1% of the adjusted variance in RPE could be explained by exercise intensity measured by heart rate alone. Whereas, Foster (1998) reported individual correlations between session RPE and heart rate load ranging from 0.75 to 0.90. Stronger relationships have also been observed between the referees' heart rates and RPE scores during fitness training sessions when compared to the relationship observed during competitive matches.

Given the strength of the relationship between match heart rate and RPE reported in Study II it may be that a dissociation was observed between these two variables during competitive matches in that the referees may have been using the RPE score as a measure of psychological as opposed to physiological stress even though the rating is psychophysiological. For example, just as blood lactate measures during soccer are influenced largely by the activity pattern that precedes the sampling, it may be that the referees' RPE ratings could have been influenced by the number of critical decisions they awarded during the game, such as yellow and red cards and also penalty kicks, even though they were instructed to use the rating as a measure of global match intensity. With this in mind future data collection may look to include two ratings; one on the technical demands of the match and separate rating on the match physical demands.

Whilst there may be limitations of both measures used to examine exercise intensity in Study II - with heart rates providing a possible overestimation of exercise intensity during competitive exercise and the referees' RPE scores possibly reflecting more a measure of technical demand as opposed to physical demand given its relative disassociation with heart rate - the use of these measures remains widespread in soccer due to the practicality of data collection and the relatively low cost when compared to other measures such as blood lactate.

The data collected in this study data formed a conference communication at the annual BASES conference² in 2004. Also, match officials' heart rates at the

² Weston, M., Helsen, W., Bird, S. & Nevill, A.M. (2005) The effect of match standard on objective and subjective match demands of elite soccer referees. *Journal of Sports Sciences*, 23, 189.

highest standard of soccer competition, the FIFA World Cup, were presented to the BASES annual conference³ in 2003 entitled 'Heart rate responses of referees and assistant referees during FIFA 2002 World Cup matches'. However, methodological issues with regards to the determination of HR_{max} make comparisons between the different competitions difficult.

The collection of match heart rate and RPE data up to this point had provided an indication of the previously unreported physical demands of English soccer refereeing. However, research presented at the 2005 BASES annual conference⁴ entitled 'The relationship between heart rate and intensity of match play in soccer referees' showed only low to moderate correlations between heart rate and match activities. Specifically, match heart rate load, which was computed by multiplying the accumulated duration in each of five different heart rate zones by a multiplier ($<60\%HR_{max} = 1$; $60-75\% = 2$; $76-85\% = 3$; $86-93\% = 4$; $>93\% = 5$) and summing the results (Edwards, 1993) demonstrated a moderate correlation with the total distance covered during a match ($r = 0.370$, $p > 0.05$, $n = 18$) and only a small correlation with the amount of high-intensity running ($r = 0.205$, $p > 0.05$, $n = 18$).

The monitoring of a referee's heart rate is much more convenient than it is for players as they are not involved in physical contacts and heart rate itself can be a useful indicator of circulatory strain (Drust, Atkinson & Reilly, 2007). Up to this point the collection of match heart rate data had provided a measure of match

³ Weston, M., Helsen, W. & DeClerk, R. (2004) Heart rate responses of referees and assistant referees during FIFA 2002 World Cup matches. *Journal of Sports Sciences*, 22 (3), 261.

⁴ Weston, M. & Castagna, C. (2005) The relationship between heart rate and intensity of match play in soccer referees. *Journal of Sports Sciences*, 23 (11/12), 1302-1303.

demands and from a practical perspective examination of heart rates together with match RPE data helped to provide valuable information as to whether the Premier League referees were meeting the physical demands of their matches. However, these results indicated that soccer referees' match heart rates alone may not represent a valid measure of match intensity, as they demonstrated low to moderate correlations with match activities. Therefore, caution should be taken when evaluating a referee's intensity of match play using heart rates alone and that match heart rates and match physical performance may represent two different measures; match heart rates reflect the physiological and psychological load placed upon the referees during actual match play, whereas high-intensity running represents an external load. Therefore, to fully determine the intensity of match play in Premier League soccer referees I suggest that both measures are used.

STUDY III

Weston, M., Castagna, C., Impellizzeri, F., Rampinini, E. & Abt, G. (2007) Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates. *Journal of Science and Medicine in Sport*, 10, 390-397.

As part of an applied research model for the sports sciences Bishop (2008) stated that once descriptive studies have been performed sports scientists then need to produce research that better understands the factors that are likely to affect performance. Therefore, with this in mind, and also in light of the findings

presented with regards to the relationship between match heart rates with match activities, Study III was a clear progression from Study II in that the referees' physical match performances (an external load), and some of the factors that may affect these loads, were analysed in detail. Study III reported a mean match total distance covered during Premier League matches of $11622 \pm 739\text{m}$ a distance higher than those reported within the more recent scientific literature (Krustrup & Bangsbo, 2001; D'Ottavio & Castagna 2001). The differences in distance covered between this study and those described in the literature may have been due to differences in the fitness of top-level referees over recent years, as the referees involved in Study III were full-time professionals, or possibly differences in the technologies used to measure the distances covered.

It has been reported that standing, walking and jogging account for more than 75% of match activity in soccer referees (Krustrup & Bangsbo, 2001). Therefore, the total distance covered during a game is a poor measure of the physical stress placed on a referee during a match (Krustrup & Bangsbo, 2001). It is the amount of high-intensity exercise that best indicates the demands of the game and in turn the development of fatigue (Bangsbo, Nørregaard & Thorsøe, 1991; Mohr, Krustrup & Bangsbo, 2003) and is therefore a more sport-specific dependent variable (Atkinson & Nevill, 2001). Of the 11622m of total distance covered by the English Premier League referees, $787 \pm 245\text{m}$ were covered during high-intensity running ($>19.8\text{km}\cdot\text{h}^{-1}$). In comparison, D'Ottavio & Castagna (2001) reported a mean high-intensity running ($>13.1\text{km}\cdot\text{h}^{-1}$) distance of $1556 \pm 493\text{m}$, while Krustrup & Bangsbo (2001) reported 1670m (range 900m – 2390m) of match high-intensity running ($>15.0\text{km}\cdot\text{h}^{-1}$) for Danish referees. Substantial

differences in the methodologies the authors used to classify high-intensity running make direct comparisons between the results of Study III and those previously reported within the literature not possible. Therefore, given these differences in high-intensity running determination it was suggested that it may be better to express this variable relative to the referees' individual maximal speed levels, providing that methodological issues regarding the determination of maximal speed and appropriate speed zones can be addressed.

Abt & Lovell (2009) recently investigated the individual applicability of the high-intensity speed zone threshold used by Prozone[®] (19.8km.h⁻¹). The authors reported that in a group of professional soccer players the second ventilatory threshold, a threshold used to delimit moderate and high-intensity zones in the prescription of training (Lucia, Hoyos, Perez & Chicarro, 2000), occurred at a mean running speed of 15.3 ± 0.7 km.h⁻¹. The difference between the median ventilatory speed and the default Prozone[®] high-intensity classification was therefore -4.8km.h⁻¹, which represented a 24% reduction in high-intensity speed and ultimately an underestimation of high-intensity running distance during a match. However, the authors did report a range of 14-16 km.h⁻¹ for the players' second ventilatory threshold, which demonstrates a variable between-subject threshold level. This in turn suggests that an arbitrary speed value for high-intensity running will not hold true on an individual basis. There are also issues with regards to the application of laboratory determined physiological thresholds in real-match scenarios. For example, Reilly (1997) reported that around 16% of the distance covered by players is in moving backwards, sideways or 'jockeying' for position and the energy cost of unorthodox modes of motion increases

disproportionately with the speed of motion (Reilly & Bowen, 1984). As Green (2006) stated ‘artificially controlled research does not fit the realities of practice’ and this may well be the case when attempting to translate a laboratory determined individual physiological threshold into a match scenario as the relationship between energy expenditure and running speed will no doubt become non-linear.

As previously mentioned, to fully determine the intensity of match play in Premier League soccer referees it is important to use both external and internal loads as this is crucial for an understanding of the distances covered and in turn how the body physiologically responds to the match load. To accurately prescribe training loads and the timing of loads following matches the sports scientist / fitness coach needs to be aware of the impact of both these measures on the individual. For example, measures of exercise intensity such as heart rate are very often expressed as values relative to the individuals HR_{max} . Therefore, it may be that the distances covered in different speed zones during matches are expressed as absolute distances covered and also as distances relative to the individual’s fitness and/or maximal running speed. Whether the speed zones used during systems such as Prozone[®] are based on an individual physiological response to an incremental field-based fitness test or as a percentage of maximal speed is an area that requires further investigation.

Ideally the determination of maximal running speed should be made using similar data collection procedures as those employed by the match analysis system. For example, the Prozone[®] system measures peak speed as the peak

value recorded during 0.5-s samples during pitch running. Therefore, to replicate this I advise that maximal running speed be determined on the field of play using data collection intervals as close as possible to one every 0.5 seconds and with the distance run being far enough to ensure top speed is attained. For example, measuring over short distances (<20m) may simply reflect acceleration speed and not peak velocity given that Young *et al.* (2008) reported common variances of only 25% and 42% between 'flying' velocity times, 20-40m and 20-30m, respectively with 10-m time. However, once an accurate maximal speed value is obtained then difficulty would lie in the determination of the individual zones. Whereas, if the speed zones were to be based on a physiological response, i.e., heart rate, blood lactate, to ideally a field-based, game specific protocol as opposed to an incremental treadmill protocol it could be possible to use zones suggested within the literature.

Catterall *et al.* (1993) reported from their findings that a referee's activity was acyclical but varied in parallel with the players' actions. However, no direct comparison was made between the players' work rates and those of the referees. Therefore, Study III reported for the first time that English soccer referees' physical match performances were partly related to the overall match intensity, as indicated by the amount of high-intensity running performed by the players (Mohr *et al.*, 2003). As a result we recommended that when assessing a referee's physical match performance there is a need to first assess the overall match intensity to gain an understanding into the pattern of the match and to assess whether or not the referee met the physical demands of the match. Consequently

all of PGMOL referees' physical match performances are now examined in context of the overall intensity at which the match they officiated was played.

A further investigation into the relationship between referee and player work rates during FA Premier League matches has been performed and presented to the 2009 BASES annual conference⁵. The referees covered significantly more total distance when compared to the players ($11326 \pm 732\text{m}$ vs. 10809 ± 386 , $P < 0.05$). There were no between-group differences for high-intensity running (737 ± 241 vs. $711 \pm 99\text{m}$, $P = 0.96$). However, the players' sprint distance was greater than that recorded by the referees ($268 \pm 79\text{m}$ vs. $161 \pm 114\text{m}$, $P < 0.05$). Correlations were observed between the referees' and players' total distance ($r=0.405$, $P < 0.001$, $n=186$), high-intensity running ($r=0.477$, $P < 0.001$, $n=186$) and sprinting ($r=0.571$, $P < 0.001$, $n=186$).

The differences between volume and intensity of running may primarily reflect the role-specific nature of referees and players during matches. Specifically, referees have to keep up with play at all times to ensure correct judgments; whereas, positional requirements of players impact upon their involvement in the game. However, these data may not illustrate the full extent of the work rate comparison between the referees and players given the use of the same absolute speed zone classifications for high-intensity running and sprinting. Also, as Abt & Lovell (2009) reported a possible underestimation of the amount of high-intensity running during competitive matches performed by players using the Prozone[®] speed thresholds, it is likely that this underestimation will be more

⁵Weston M., Portas, M., Barnes C., Harley, J. & Bland, M. (2009) A comparison of FA Premier League referee and player match work rates. *Journal of Sports Sciences*

pronounced for the referees, given their age, training status and ultimately fitness levels in comparison to professional soccer players.

The relationship between the referees' work rates with those of the players was stronger than previously reported in Study III, especially with regards to sprinting. This further reinforces PGMOL practice of evaluating the overall intensity of the match prior to examining the referees' work rate profile to ensure that the referees are meeting the physical demands of their matches.

To obtain a more in-depth assessment of the relationship between referees' work rates with those of the players involved in the same game I recommend that a more intricate analysis should be performed. For example, Mohr *et al.* (2003) showed the importance of analysing and reporting data with sufficient magnification to enable small segments of matches (5-min periods) to be compared. Drust *et al.* (2007) stated that this type of approach would enable the intricacies of the work rate pattern and the factors that affect it to be thoroughly understood. This way changes in a referees' work rate pattern across a game can be evaluated against changes observed for the players, as opposed to comparing global match work rates. Also, given that the players' match activities are related to their playing position (DiSalvo, Gregson, Atkinson, Tordoff & Drust, 2009) a comparison of the referees' work rates with the different playing positions would provide a further examination of the relationship between referee and player activity profiles.

A limitation of the analysis between the referee and player work rates performed in Study III was that comparisons between referee and player work rates only considered the players who completed the full match. Given that substitutes perform more high-intensity running than the other players towards the end of the match (Mohr *et al.*, 2003) this methodology may not have accurately reflected match intensity across the whole match. Consequently, I propose that any future comparisons between referee and player work rates involve the total distances covered by each individual player on the field at the time during each 5-min period. This process will permit the overall match intensity to be evaluated as it would include data from substitutes.

A further limitation of this study was that given the strength of our results ($r = 0.43$) the physical match performances of the referees were probably influenced by other factors not measured within the present study. Therefore, a more complex analysis of the relationship between referees' physical match performances in relation to those of the players should be performed, i.e., league position of teams involved, styles of play, effect of half-time score line on second half match activities.

Stølen, Chamari, Castagna & Wisløff (2005) reported that the analysis of between-half distance coverage is of great interest as it can reveal the occurrence of fatigue and/or refereeing strategies; both of which would be classified as determinants of overall physical match performance. Through the use of the median-split technique it was possible to gain a more in-depth assessment of between-half work rates than those previously reported (Johnston &

McNaughton, 1994; Krstrup & Bangsbo, 2001; D'Ottavio & Castagna, 2001). Using this type of analysis the results demonstrated that a higher amount of total distance and high-intensity running completed during the first half resulted in a lower total distance and high-intensity running performed by the referees in the second half. Whereas, the lower the amount of total distance and high-intensity running in the first half, the greater the amount of total distance and high-intensity running covered in the second half. Whilst statistically significant the practical relevance of the changes in total distance (2%) is questionable. However, the changes in high-intensity running is of interest as it may indicate fatigue on behalf of the referees whose high-intensity running reduced by 11% in the second half and pacing strategies from those referees who high-intensity running increased by 30% in the second half. This 'sparing behaviour' may be adopted by the referees to avoid fatigue (Castagna & Abt, 2003) and should of concern to the sports scientist as it may reflect a lack of confidence on behalf of the referees with regards to their fitness levels.

Alternatively, given the research design employed in this study for examining the change in between-half work rates it may be that the results were influenced by the 'regression to the mean' phenomenon. The regression to the mean is a chance finding masquerading as a real one, in that when a sample of subjects is investigated in a test-retest fashion it will be extremely likely that the subjects who score higher than average values on the initial test will show lower values, closer to the sample mean (Atkinson, Waterhouse, Reilly & Edwards, 2001). Conversely, subjects who score particularly low values on the initial test will be likely to show an increase on the second test, and this increase is purely a

statistical artefact (Atkinson et al., 2001). Therefore future analysis should correlate any changes with the average match distances.

Ultimately, given that the regression to the mean is a problem where the intra-individual standard deviation is large relative to the inter-individual standard deviation (Shephard, 2003) and that within-subject coefficients of variation of 15 and 47% were reported in Study III for total distance and high-intensity running, respectively, it maybe that future studies analysing match activity data should avoid research designs that are susceptible to the regression to the mean. Shephard (2003) reported that regression to the mean can be eliminated when using a fully controlled experimental design and that if interesting trends are suggested by an inspection of the results, these trends are best verified by well-considered new research.

Match to match variations within a 7-day period demonstrated that referees' high-intensity running was significantly reduced during the second match. These findings have real-world implications for the bodies who appoint referees to matches as they demonstrate that appropriate planning of a referee's fixtures is required to enable the referees to recover sufficiently and in turn be able to cope with the physical demands of all of their matches. Once again the implications for the development of appropriate fitness levels are apparent from such findings. However, these data were not examined in the context of the overall match intensity as determined by the players' high-intensity running.

In Study III a reduced match high-intensity running may have represented fatigue on behalf of the referee or merely a slower tempo of game. As previously mentioned a more detailed analysis of match activity profiles would enable the intricacies of the work rate pattern to be more thoroughly understood. For example, Mohr *et al.* (2003) reported that high-intensity running in players during the 5-min period that immediately followed the peak 5-min high-intensity running period was below the match 5-min average distance. The authors attributed this decrease to temporary fatigue. Mallo, Navarro, Garcia-Aranda & Helsen (2008) also reported this phenomenon to occur in elite-standard soccer referees, suggesting that referees can be exposed to this kind of temporary fatigue at different stages of the match.

Therefore, when attempting to evaluate fatigue in soccer referees during matches activity profiles should not only be examined across the match but also immediately after the referees' most intense periods of running. However, when undertaking this type of analysis it is recommended that each individual referee work rate period is examined alongside the changes in work rate observed for the players during that particular period and possibly even the amount of time the ball was in play. I recommend that analyses of referees' physical match performances should consistently be performed in-line with the contemporary research on match activity patterns in soccer and given that the most important aspect of refereeing is the decision-making process (Helsen & Bultynck, 2004) a referee's match activity profile should ultimately be examined in relation to their accuracy of decision making, providing a reliable and valid index of 'correctness' can be determined.

The analysis of the referees' between-half work rates was presented at the 6th World Congress of Science & Football⁶, held in Turkey 2007. Also presented was an invited lecture entitled 'Full-time referees; the role of the Sports Scientist'. A similar presentation was given at the Association des Chercheurs en Activités Physiques et Sportives (ACAPS) conference in Leuven, Belgium, 2007 to a Union des Associations Européennes de Football (UEFA) sponsored soccer referee symposium.

6.3 AGEING AND PHYSICAL PERFORMANCE IN SOCCER

REFEREES

STUDY IV

Castagna, C., Abt, G., D'Ottavio, S. & Weston, M. (2005) Age related effects on fitness performance in elite-level soccer referees. *Journal of Strength and Conditioning Research*, 19, 785-790.

STUDY V

Weston, M., Castagna, C., Impellizzeri, F. et al. (2010) Ageing and Physical Match Performance in English Premier League Soccer Referees. *Journal of Science and Medicine in Sport*, 13 (1), 96-100.

Given the age difference that exists between referees and players, and that ageing has a negative impact upon fitness levels (Stamford, 1988) the aims of studies IV and V were to gain an insight into 1) the implications of increasing age on

⁶ Weston, M., Castagna, C., Impellizzeri, F., Rampinini, E. & Abt, G. (2007) Analysis of between-half work rates in English Premier League soccer referees. *Journal of Sports Science and Medicine*, S10, 210.

referees' fitness levels, and 2) the impact of ageing upon referees' physical match performances during competitive matches.

Study IV reported no significant difference in endurance performance between age groups when assessed as distance covered during the referees' fitness test (12-min run), even though regression analysis revealed the existence of a significant trend towards an age-related reduction in performance ($r=-0.52$). The absence of any between-group effects may have been a consequence of a reduced statistical power resulting from the categorisation of continuous variables (Altman and Royston, 2006). When a further analysis was performed in the laboratory the young (33.1 ± 1.8 years) referees had higher VO_2max levels when compared to the older (42.1 ± 1.2 years) referees, although the older referees were able to match the young group for running economy at a speed of $10\text{km}\cdot\text{h}^{-1}$.

All referees involved in the study were given a two-month training programme to follow in the lead up to testing. The programme consisted of continuous running, high-intensity running and sprinting and there were no differences in recorded training intensity across the three sub-groups; young, intermediate and old. Training intensity was regulated on an individual basis using heart rate monitoring; no running speeds were monitored. It has been previously reported that soccer referees training during the season used to consist of continuous running at intensities $\sim 75\%\text{HR}_{\text{max}}$ over 3 – 7km (Krustrup & Bangsbo, 2001). This may explain why the older referees, despite lower VO_2max values, were able to match the young referees with regards to running economy as the training programme was only short-term in duration and athletes' most economical

velocities or power outputs tend to be those at which they habitually train (Jones & Carter, 2000). It may also account for why the referees involved in the present study showed age-related differences in the short (1-35 seconds), predominantly anaerobic domain and specifically in the explosive strength and anaerobic endurance tests. The practical relevance of such a finding can be seen in the need for sports scientists and fitness trainers to include specific anaerobic conditioning drills within the training programmes of referees, especially as age progresses.

Whilst a valuable insight into the effect of ageing and physical fitness was obtained for soccer referees in Study IV, a major limitation of the study was that the relationship between age and physical match performance was not reported. Another important issue with regards to ageing and soccer referees is that international refereeing governing bodies have an enforced retirement age of 45 years. However, in some countries this enforced retirement has been successfully challenged in a court of law as it is considered to be against the European Employment Directive, which prohibits age discrimination in employment. Consequently, the aim of Study V was to examine the effect of age upon the physical match performances and match physiological loads of English Premier League soccer referees.

Study V demonstrated for the first time a clear age-related decline in elite-level soccer referees' physical match performances as the total distance covered, high-intensity running and the number of sprints performed all decreased as referee age increased. This occurred despite there being no between-group differences in overall match intensity, as determined by the amount of high-intensity running

performed by the players. Therefore, part of the methodology employed in Study V was a direct consequence of findings reported in Study III, thus showing how earlier work helped to inform future research practice.

The decline in physical match performance may have been related to the age associated impairments in physical capacities already reported in soccer referees. Furthermore, the fact that the older referees perceived their match demands to be higher, despite covering less ground and at the same heart rate, the higher RPE might suggest greater fatigue. In this regard the examination of the relationship between match heart rates and RPE scores (internal load) with the actual distances covered (external load) by the referees was a clear progression from the research designs presented in Studies II and III.

Alternatively, given that the literature examining the role of increasing age on fitness has examined an age range older than the referees involved in Studies IV and V it may be that the age range employed in the current study (31-48 years) was not great enough to elicit substantial reductions in aerobic fitness that would impair the referees' match activity profiles. For example, in a study examining the age-related changes in marathon and half-marathon performances Leyk *et al.* (2007) reported a main finding that in a sample of trained subjects significant age-related losses in endurance performance did not occur before the age of 50 years. Whilst age associated decrements in physical match performance were observed in Study V, no significant relationship was observed between age and the referees' average distance from the ball and, more importantly, from fouls. Therefore, despite covering less total distance, high-intensity running and

performing fewer sprints the older referees (43-48 years) were able to maintain an average distance from fouls that was comparable to that recorded by the young referees (31-36 years). As a result, it maybe the effectiveness of the referees' match activities in ensuring the best possible viewing positions that is of most relevance, which is more likely to be a component of experience given that experience leads to better anticipatory responses (Abernethy, 1987). Consequently it is possible that the independent variable examined in Study V was one of referee experience as opposed to referee age.

From 2003 to 2008 the maximum age range for the English Premier League referees was no greater than 17 years. However, with the abolishment of age imposed retirement combined with the promotion of younger referees this range has increased to 25 years. Experience is considered a prerequisite to officiate at elite-level by national and international refereeing bodies yet the age at which referees are now being promoted is reducing. This will no doubt impact upon the experience of the referees at elite-level as MacMahon, Helsen, Starkes & Weston (2007) reported that actual match performance is a significant activity for skill acquisition and refinement. Consequently refereeing governing bodies should attempt to address this potential problem via appropriate talent identification programmes and the possible establishment of refereeing academies.

The findings reported in Study V should clearly influence the practice of referee retirement as these data would appear to suggest that an enforced retirement age of 45 years may not be justified by the inability of the referees to keep up with the play. Or, it may be that the enforced retirement age of 45 years has become

dated with the increased fitness levels previously reported in referees following more modern training regimes (Study I) and also the advent of full-time, professional soccer referees. Therefore I recommend that a referee's retirement should be undertaken considering ability to officiate alongside physical performance during fitness tests and training sessions and also match activity profiles; a practice now employed by the PGMOL. For example, when a match official applies for an extension beyond the previous retirement age a detailed report on each referee is prepared and this covers performance on fitness tests, performances in training and the match officials' physical match performances. Where possible, data should be reported over multiple seasons so that any longitudinal changes can be highlighted.

6.4 FITNESS AND MATCH PERFORMANCE

STUDY VI

Weston, M., Castagna, C., Impellizzeri, F. & Helsen, W. (2009) Relationship among field-test measures and physical match performance in elite-standard soccer referees. *Journal of Sports Sciences*, 27 (11), 1177-84

A limitation of study V was that it was not possible to report on the referees' fitness levels as determined through regular fitness tests even though a clear decrement in physical match performance was observed. Study III had also reported on the referees' match activities yet not placed them in context with the referees' fitness levels. Therefore, in comparison to these studies, strength of

Study VI was that for the first time in this thesis referees' fitness **levels** and physical match performances were examined together, with the study's aim being to examine the validity of the referees' fitness tests as a measure of match-related fitness.

Each close-season referees are required to pass a fitness test before they can take up appointments. Until recently referees were required to pass a battery of field-based fitness tests, set by FIFA that included 2x50 m and 2x200 m sprints, followed by the 12-min Cooper test (Eissmann & D'Hooghe, 1996). However, research has illustrated that these tests are poor measures of a referee's match-related physical capacity. Castagna, Abt & D'Ottavio (2002) reported that the 12-min run showed large and moderate correlations with high-intensity running and distance covered at maximal speeds ($r=0.51$ and $r=0.32$, respectively). However, from a practical perspective the strength of these correlations needs to be higher in order to support construct validity as Implellizzeri & Marcora (2009) stated that a correlation of larger than 0.70 between a new test and the reference measure is conventionally used as a benchmark for construct validity. No relationship was evident between the sprint tests with distance covered at high-intensity running and sprinting (Castagna *et al.*, 2002).

To overcome this lack of fitness test validity, FIFA recently replaced these fitness tests with a new battery of field-based fitness tests, reported to be more of a measure of a referee's match-related physical capacity (FIFA, 2008). The test battery involves an aerobic measure (a high-intensity 150m interval test) and a sprint / repeated-sprint measure (6x 40m sprints). FIFA established minimum

selection criteria in that should a referee attain these standards their fitness is deemed sufficient to be able to cope with match demands.

The referees' performances on the interval run, as determined by their heart rate response, demonstrated a large correlation (-0.70) with total distance covered during match. A large correlation was also observed between the referees' interval run heart rate response and match high-intensity running (-0.57). In terms of the referees' sprint performance, both the fastest and average 40m sprints demonstrated a very large correlation with match high-intensity running (-0.76; -0.77) and also match sprinting distance (-0.75; -0.77). However, a field test cannot be considered sport-specific until a direct association between the most relevant aspects of physical match performance and the field test investigated has been determined (Castagna, Impellizzeri, Rampinini, D'Ottavio & Manzi, 2008). Therefore, it was concluded that with the current test protocol as it stands, only the 40m repeated-sprint test showed appropriate construct validity to be used as an assessment of a referees' match-related physical capacity.

The median-split technique was used in Study VI in an attempt to quantitatively demonstrate that those referees who performed best on the fitness tests produced the best physical match performances. Whilst these results supported the correlation analyses it is recommended that future studies avoid such an analysis technique as dichotomising continuous variables leads to a reduction in statistical power and also characterises individuals close to but on opposite sides of the cutpoint as being very different rather than very similar (Altman and Royston,

2006). Altman and Royston (2006) went on to recommend that instead of categorising continuous variables they should be kept continuous as then a linear regression could be used.

Performance in most sports is the result of a blend of several factors and the sports scientist can, through physiological testing of the participants, analyse these factors and use the information to provide individual profiles of their respective strengths and weaknesses (Svensson & Drust, 2005). Despite the **low** validity, both logical and construct, of the interval run, FIFA still employ this fitness test, along with the repeated-sprint test as a referee selection criterion at the beginning of each season and also at major tournaments. Therefore, as it stands this protocol has to be carried out on all the Premier League referees prior to each season.

Initially the FIFA fitness tests were also run during the season as an evaluation of the referees' fitness levels and ultimately the effectiveness of the training programmes supplied by sports scientists as objective data are required on changes in performance over time to study the effectiveness of a training programme (Balsom, 1994). YYIR1 performance has been reported to be more closely related to the physical performance of top-class referees during a soccer match (Krustrup & Bangsbo, 2001), when compared to the FIFA interval run. Consequently, given the strength of the results presented by Krustrup & Bangsbo (2001) against the strength of those reported for the interval run in Study VI the YYIR1 is now used as the referees' in-season fitness test. This practice is also reinforced by the observation that the YYIR1 is a valid in-season tool to test

football specific endurance (Castagna, Impellizzeri, Chamari, Carlomango & Rampinini, 2006; Krstrup *et al.*, 2003) and Study III reported that the referees' physical match performances are, in part, related to the physical performances of the players.

It is recommended that given the strength of the results in Study VI if FIFA want their fitness tests to accurately reflect a match official's match-related physical capacity then the use of the YYIR1 should be considered. This test should be performed alongside a repeated-sprint test that has a protocol more in-line with previously validated tests such as Impellizzeri *et al.* (2008). Furthermore, since all the referees in Study VI reached the minimum level set by FIFA to officiate at international level, yet produced wide-ranging physical match performances, the ability of the current FIFA tests for the selection of referees based on physical performance should be further investigated by purposely made studies. For example, if the aim of a test is to select athletes, it should be able to discriminate individuals of different competitive levels (Impellizzeri & Marcora, 2009). Yet to account for different standards of refereeing (national vs. international), FIFA modified the test procedures by setting a slower minimum standard on the repeated-sprint test (6.2 seconds) and an increased recovery interval duration on the interval run (40 seconds) for national standard referees. Whilst this approach may be commended given that the intensity of match play in referees was demonstrated in Study II to increase with an increase in competition standard, the discriminate ability of these tests to distinguish between different competitive levels has clearly not been addressed by FIFA. Methods such as the receiver operator characteristics (ROC) curve can be used to validate the discriminant

ability of a performance test (Implellizzeri & Marcora, 2009). However, such a method requires the setting of a criterion score and the setting of a FIFA minimum standard would still require empirical validation, which raises the issue of how much total distance, high-intensity running and sprinting should be used as the reference measure as being 'fit for purpose'.

Consequently, further research is needed to establish a criterion distance for total distance, high-intensity running and sprinting that would enable referees to adequately meet the physical demands of their matches. This could be achieved via a series of comparisons between physical match performances against technical performances. For example, national and international refereeing bodies have criteria established for technical marks (awarded by a match assessor) and the referees' physical match performances could then be compared against these scores. However, the marks awarded by the assessor are likely to be affected by intra- and inter-observer variability.

The high match-to-match variability that has been observed in referees' and players' physical match performances raises another potential problem concerning the setting of a criterion-related fitness test minimum standard. Recent data analysis has shown between-subject coefficients of variation for total distance, high-intensity running and sprinting of 3.6%, 14.2% and 31.9%, respectively for Premier League players on a total of 237 matches during the 2008/09 English football season. These values were significantly higher for the referees (6.4%, 35.4% and 87.9%), thus demonstrating a more heterogeneous group. Such variability has serious implications for sample size estimation and

may ultimately suggest that previous studies examining referee physical match performance pre- and post-intervention may be underpowered. For example, using a nomogram to estimate the effects of measurement repeatability error Batterham & Atkinson (2005) reported that to detect a meaningful difference of 10% using a measure with a coefficient of variation of 30% would require a sample size of approximately 200 games. Whereas, when examining the impact of intense intermittent exercise training on physical match performance in a group of elite-level soccer referees Krstrup & Bangsbo (2001) examined only eight referees on one competitive game each post intervention. Therefore, further work is required to fully determine the extent of match-to-match variability of physical match performance variables in soccer referees and also the factors that may impact upon this variability that is, experience, age, standard of competition, time of season.

One final issue regarding the use of a minimum standard on referees' fitness tests is that this may lead to referees not producing maximal performances, in that once the standard has been achieved they withdraw from the test. This may be of particular concern when fitness tests are undertaken during the season and the referee's focus is on match preparation and not a maximal fitness test. This problem may be unavoidable given that the referees are not monitored in training on a daily basis as per professional players, and whilst heart rates and ratings of perceived exertion scores do provide valuable information as to the referees' intensity of training they do not reflect the external loads placed on the referees (i.e., distances covered and at what speeds, unless track or treadmill running is performed).

More recently global positioning satellite (GPS) systems have been used to assess the physical demands of team sports during competition and training (Edgecomb & Norton, 2006). GPS units are now commercially available and some of the devices incorporate heart rate telemetry and accelerometers, both of which can provide additional measures of physiological strain and training load (Macleod, Morris, Nevill & Sunderland, 2009). Therefore, it may be that the use of these devices during training and also matches would enable the referees' external and internal loads to be analysed in more detail, which in turn would enable more individual specific training loads and recovery protocols to be developed.

In conclusion, this programme of work examined some of the contemporary issues regarding the applied physiology of soccer refereeing. Each study attempted to use an appropriate design that best matched research method to problem. In doing this the programme of work did not remain at a fixed point on the basic-applied research continuum but shifted with the experimental need of each individual study. Therefore, it is hoped that the studies constituting this PhD submission demonstrate it is possible to exist somewhere within a 'middle ground' bearing in mind that the strengths of one type of research are the weaknesses of the other and vice versa.

References

- Abernethy, B. (1987) Anticipation in sport. *Physical Education Review*, 10, 5-16.
- Abt, G. & Lovell, R. (2009) The use of individualized speed and intensity thresholds for determining the distance run at high-intensity in professional soccer. *Journal of Sports Sciences*, 27, 893-898
- Altman, D.G. & Royston, P. (2006) The cost of dichotomising continuous variables. *British Medical Journal*, 332, 1080.
- Atkinson, G. & Nevill, A.M. (2001) Select issues in the design and analysis of sports performance research. *Journal of Sports Sciences*, 19, 811-827.
- Atkinson, G., Waterhouse, J., Reilly, T. & Edwards, B. (2001) How to show that unicorn milk is a chronobiotic: the regression-to-the-mean statistical artefact. *Chronobiology International*, 18, 1041-1053.
- Balsom, P.D. (1994) Evaluation of physical performance. In B. Ekblom (Ed.), *Football (soccer)*, pp. 102-123. London:Blackwell.
- Bangsbo, J., Nørregaard, L. & Thorsøe, F. (1991) Activity profile of competition soccer. *International Journal of Sports Medicine*, 13, 125-132.
- Batterham, A.M. & Atkinson, G. (2005) How big does my sample need to be? A primer on the murky world of sample size estimation. *Physical Therapy*, 6, 153-163.
- Bishop, D. (2008) An applied research model for the sports sciences. *Sports Medicine*, 38, 253-263.
- Castagna, C. & Abt, G. (2003) Intermatch variation of match activity in elite Italian soccer referees. *Journal of Strength and Conditioning Research*, 17, 388-392.

- Castagna, C., Abt, G. & D'Ottavio, S. (2002). Relationship between fitness tests and match performance in elite Italian soccer referees. *Journal of Strength and Conditioning Research*, 16, 231-235.
- Castagna, C., Impellizzeri, F.M, Chamari, K., Carlomango, D. & Rampinini, E (2006). Aerobic fitness and Yo-Yo continuous and intermittent test performances in soccer players: a correlation study. *Journal of Strength and Conditioning Research*, 20, 320-325.
- Castagna, C., Impellizzeri, F.M., Rampinini, E., D'Ottavio, S. & Manzi, V. (2008). The Yo-Yo intermittent recovery test in basketball players. *Journal of Science and Medicine in Sport*, 11, 202-208.
- Catterall, C., Reilly, T., Atkinson G. & Coldwells, A. (1993) Analysis of the work rates and heart rates of association football referees. *British Journal of Sports Medicine*, 27, 193-196.
- Coutts, A., Rampinini, E., Marcora, S.M., Castagna, C. & Impellizzeri, F.M. (2007) Heart rate and blood lactate correlates of perceived exertion during small-sided soccer games. *Journal of Science and Medicine in Sport*, 12, 79-84.
- Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P. & Drust, B. (2009) Analysis of high intensity activity in Premier League soccer. *International Journal of Sports Medicine*, 30, 205-212
- D'Ottavio, S. & Castagna, C. (2001) Physiological load imposed on elite soccer referees during actual match play. *Journal of Sports Medicine and Physical Fitness*, 41, 27-32.
- Drust, B., Atkinson, G. & Reilly, T. (2007) Future perspectives in the evaluation of the physiological demands of soccer. *Sports Medicine*, 37, 783-805.

- Edgecomb, S.J. & Norton, K.I. (2006) Comparison of global positioning and computer-based tracking systems for measuring player movement distance during Australian football. *Journal of Science and Medicine in Sport*, 9, 25-32.
- Edwards, S. (1993) High performance training and racing. In *The Heart Rate Monitor Book*. (Edited by S Edwards), pp 113-123. Sacramento, CA: Feet Fleet Press.
- Eissmann, H.J. & D'Hooghe, M.D. (1996). Sports medical examinations. In *The 23rd man: Sport medical advice for football referees*. (Edited by H.J. Eissmann), pp. 7-19. Leipzig: Gersöne-Druck.
- FIFA (Fédération Internationale de Football Association). (2008). *Refereeing 2008: International lists*, pp. 246-249. Zurich: FIFA.
- Foster, C. (1998) Monitoring training in athletes with reference to overtraining syndrome. *Medicine and Science in Sports and Exercises*, 30, 1164-68.
- Gellish, R.L., Goslin, B.R., Olsen, R.E., Macdonald, A., Russi, G.D. & Mougdil, V.K. (2007) Longitudinal modelling of the relationship between age and maximal heart rate. *Medicine and Science in Sports and Exercise*, 27, 822-829.
- Green, L.W. (2006) Public health asks of systems sciences: to advance our evidence-based practice, can you help us get more practice-based evidence? *American Journal of Public Health*, 96, 406-409.
- Helgerud, J., Engen, L.C., Wisløff, U. & Hoff, J. (2001) Aerobic training improves soccer performance. *Medicine and Science in Sports and Exercise*, 33, 1925-1931
- Helsen, W. & Bultynck, J.B. (2004) Physical and perceptual-cognitive demands of top-class refereeing in association football. *Journal of Sports Sciences*, 22, 179-189.

- Impellizzeri, F.M. & Marcora, S.M. (2009) Test validation in sport physiology: lessons learned from clinimetrics. *International Journal of Sports Physiology and Performance*, 4, 269-277.
- Impellizzeri, F.M., Rampinini, E., Castagna, C., Bishop, D., Ferrari Bravo, D., Tibaudi, A. *et al.* (2008) Validity of a repeated-sprint test for football. *International Journal of Sports Medicine*, 28, 228-235.
- Johnston, L. & McNaughton, L. (1994) The physiological requirements of soccer refereeing. *Australian Journal of Science and Medicine in Sport*, 26, 67-72.
- Jones, A.M. & Carter, H. (2000) The effect of endurance training on parameters of aerobic fitness. *Sports Medicine*, 29, 373-386.
- Leyk, D., Erley, O., Ridder, D., Leurs, M., Ruther, T., Wunderlich, M. *et al.* (2007) Age-related changes in marathon and half-marathon performances. *International Journal of Sports Medicine*, 28, 513-517.
- Lucia, A., Hoyos, J., Perez, M. & Chicarro, J.L. (2000) Heart rate and performance parameters in elite cyclists: A longitudinal study. *Medicine and Science in Sports and Exercise*, 32, 1777-1782.
- Krustrup, P. & Bangsbo, J. (2001) Physiological demands of top-class soccer refereeing in relation to physical capacity: effect of intense intermittent exercise training. *Journal of Sports Sciences*, 19, 881-891.
- Krustrup, P., Mohr, M., Amstrup, T., Rysgaard, T., Johansen, J., Steensberg, A. *et al.* (2003) The Yo-Yo intermittent recovery test: physiological response, reliability and validity. *Medicine and Science in Sports and Exercise*, 35, 697-705.

- Macleod, H., Morris, J., Nevill, A. & Sunderland, C. (2009) The validity of a non-differential global positioning system for assessing player movement patterns in field hockey. *Journal of Sports Sciences*, 27, 121-128.
- MacMahon, C., Helsen, W.F., Starkes, J.L. & Weston, M. (2007) Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, 25, 65-78.
- Mallo, J., Navarro, E., Garcia-Aranda, J.M. & Helsen, W.F. (2008) Activity profile of top-class association football referees in relation to fitness-test performance and match standard. *Journal of Sports Sciences*, 27, 9-17.
- Mohr, M., Krstrup, P. & Bangsbo, J. (2003) Match performance of high-standard soccer players with special reference to development of fatigue. *Journal of Sports Sciences*, 21, 519-528.
- Reilly, T. (1997) Energetics of high-intensity exercise (soccer) with particular reference to fatigue. *Journal of Sports Sciences*, 15, 257-263.
- Reilly, T. & Bowen, T. (1984) Exertional costs of unorthodox modes of motion. *Perceptual Motor Skills*, 58, 49-50.
- Shephard, R.J. (2003) Regression to the mean: a threat to exercise science? *Sports Medicine*, 33, 575-584.
- Stamford, B.A. (1988) Exercise and the Elderly. *Exercise and Sports Science Reviews*, 16, 341-79.
- Stølen, T., Chamari, K., Castagna, C. & Wisløff, U. (2005) Physiology of soccer: an update. *Sports Medicine*, 35, 501-536.
- Svensson, M. & Drust, B. (2005) Testing soccer players. *Journal of Sports Sciences*, 23, 601-618.

Young, W., Russell, A., Burge, P., Clarke, A., Cormack, S. & Stewart, G. (2008)
The use of sprint tests for assessment of speed qualities if elite Australian Rules
footballers. *International Journal of Sports Physiology and Performance*, 3, 199-
206.

APPENDIX

Collaborative letters of support

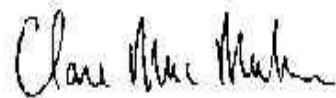
Weston, M., Helsen, W., MacMahon, C. & Kirkendall, D. (2004) The impact of specific high intensity training sessions on football referees' fitness levels. *American Journal of Sports Medicine*, 32, S54-S61.

This is to declare that, as first author of the above scientific paper, Matthew Weston:

- was involved in the formulation of the research study's rationale and aims
- helped with the data collection process
- compiled the data and performed the statistical analysis with particular cooperation from Werner Helsen
- wrote up the manuscript, taking into account comments and suggestions from the co-authors
- took the lead with regards to the revision of the manuscript as per the reviewers' comments.



.....
Werner Helsen



.....
Clare MacMahon



.....
Signature of Candidate



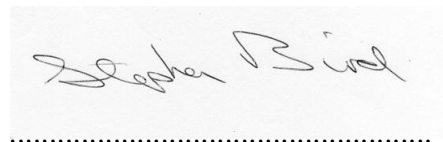
.....
Don Kirkendall

Signature of co-author (s)

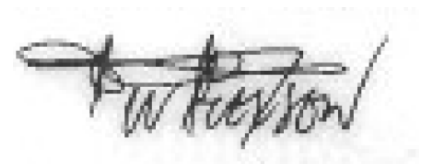
Weston, M., Bird, S., Helsen, W., Nevill, A. & Castagna, C. (2006) The effect of match standard and referee experience upon the objective and subjective match workload of English Premier League referees. *Journal of Science and Medicine in Sport*, 9, 256-262.

This is to declare that, as first author of the above scientific paper, Matthew Weston:

- initiated the study
- formulated the study's experimental design and aims with guidance from Steve Bird & Werner Helsen
- collected and compiled the data
- performed the statistical analysis with guidance from the co-authors, in particular Werner Helsen & Alan Nevill
- wrote up the manuscript, taking into account comments and suggestions from the co-authors
- took the lead with regards to the revision of the manuscript as per the reviewers' comments.



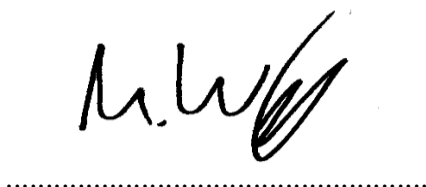
Steve Bird



Werner Helsen



Alan Nevill



Signature of Candidate



Carlo Castagna

Signature of co-author (s)

Weston, M., Castagna, C., Impellizzeri, F., Rampinini, E. & Abt, G. (2007) Analysis of physical match performance in English Premier League soccer referees with particular reference to first half and player work rates. *Journal of Science and Medicine in Sport*, 10 (6), 390-397.

This is to declare that, as first author of the above scientific paper, Matthew Weston:

- initiated the study
- formulated the studies experimental design and aims
- collected and compiled the data
- performed the statistical analysis with cooperation from the co-authors, in particular Carlo Castagna & Franco Impellizzeri
- wrote up the manuscript, taking into account comments and suggestions from the co-authors
- took the lead with regards to the revision of the manuscript as per the reviewers' comments.



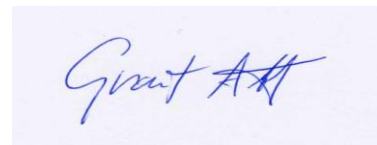
.....
Carlo Castagna



.....
Franco Impellizzeri



.....
Ermanno Rampinini



.....
Grant Abt



.....
Signature of Candidate

.....
Signature of co-author (s)

Castagna, C., Abt, G., D'Ottavio, S. & Weston, M. (2005) Age related effects on fitness performance in elite-level soccer referees. *Journal of Strength and Conditioning Research*, 19 (4), 785-790.

This is to declare that, as co-author of the above scientific paper, Matthew Weston:

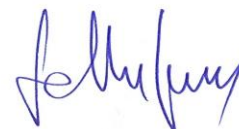
- was involved in the formulation of the research study's rationale and aims
- contributed to the data analysis, structure and writing of the manuscript
- helped with regards to the revision of the manuscript as per the reviewers' comments.



.....
Carlo Castagna



.....
Grant Abt



.....
Stefano D'Ottavio



.....
Signature of Candidate

.....
Signature of co-author (s)

Weston, M., Castagna, C., Impellizzeri, F., Rampinini, E. & Breivik, S. (2010) Ageing and Physical Match Performance in English Premier League Soccer Referees. *Journal of Science and Medicine in Sport*, 13, 96-100.

This is to declare that, as first author of the above scientific paper, Matthew Weston:

- initiated the project
- formulated the study's experimental design and aims
- collected and compiled the data
- performed the statistical analysis with cooperation from the co-authors, in particular Carlo Castagna & Franco Impellizzeri
- wrote up the manuscript, taking into account comments and suggestions from the co-authors
- took the lead with regards to the revision of the manuscript as per the reviewers comments.



.....
Carlo Castagna



.....
Franco Impellizzeri



.....
Ermanno Rampinini



.....
Signature of Candidate



.....
Simon Breivik

Signature of co-author (s)

Weston, M, Castagna, C., Impellizzeri, F. & Helsen, W. (2009) Relationship between field tests and physical match performance in elite soccer referees. *Journal of Sports Sciences*, 27 (11), 1177-1184.

This is to declare that, as first author of the above scientific paper, Matthew Weston:

- initiated the project
- formulated the study's experimental design and aims
- collected and compiled the data
- performed the statistical analysis
- wrote up the manuscript, taking into account comments and suggestions from the co-authors
- took the lead with regards to the revision of the manuscript as per the reviewers comments.



.....
Signature of Candidate



.....
Carlo Castagna



.....
Franco Impellizzeri



.....
Werner Helsen

Signature of co-author (s)